

Schnelles Testverfahren zur Evaluierung von EEPROM-Eigenschaften

Von der Fakultät für Ingenieurwissenschaften
Abteilung Elektrotechnik und Informationstechnik
der Universität Duisburg-Essen

zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften (Dr. -Ing.)

genehmigte Dissertation

von
Andre Uhlemann
aus
Lüdenscheid

1. Gutachter: Prof. Dr. -Ing. Holger Vogt
2. Gutachter: Prof. Dr. -Ing. Horst Fiedler
Tag der mündlichen Prüfung: 27.05.2011

Danksagung

Diese Arbeit ist während meiner Tätigkeit bei der ELMOS Semiconductor AG und als wissenschaftlicher Mitarbeiter am Fraunhoferinstitut für Mikroelektronische Schaltungen und Systeme entstanden. Der Entwicklungsprozess der Doktorarbeit, angefangen von der Themenfindung bis hin zur experimentellen Umsetzung und Datenanalyse, vollzog sich in einem langwierigen Prozess, in dem ich von zahlreichen Seiten durch konstruktive Diskussion unterstützt worden bin.

Für die sehr gute Betreuung bedanke ich mich bei Prof. Dr. H. Vogt

Für die sehr gute Zusammenarbeit bei den Veröffentlichungen und für die konstruktiven wissenschaftlichen Gespräche bedanke ich mich bei Dipl.-Ing. A. Aal.

Ich bedanke mich bei der NVM-Gruppe der ELMOS AG (Dipl.-Ing. T. Kleining, Dipl. Phys. J. Möller und Dipl.-Ing. K. Malina) für die gute Arbeitsatmosphäre und Hilfsbereitschaft.

Weiterhin bedanke ich mich bei meinem Vater Klaus, der mich während meiner gesamten Promotion unterstützt hat.

Am Ende bedanke mich bei allen Mitarbeitern der ELMOS Semiconductor AG und des Fraunhofer IMS, die am Gelingen der Arbeit beteiligt waren und namentlich nicht erwähnt worden sind.

Veröffentlichungen

A. Uhlemann, A. Aal, H. Vogt „A fast WLR Test for the evaluation of EEPROM Endurance“ *IEEE-IRW Final Report, 2009*

A. Uhlemann, A. Aal, H. Vogt „A fast WLR Test for the evaluation of EEPROM Endurance“ *ITG 8.5.6 Fachtagung, Dresden 2009*

Physikalische Größen

Phys. Größe	Einheit	Erläuterung
$h = 6.62 \times 10^{-34}$	Js	Planck-Konstante
$e = 1.602 \times 10^{-19}$	C	Elementarladung
$\epsilon_0 = 8.85 \times 10^{-12}$	$AsV^{-1}m^{-1}$	Influenzkonstante des Vakuums
$\epsilon_r = 3.9$	1	Dielektrizitätszahl für Silizium
$m_e = 9.10 \times 10^{-31}$	kg	Elektronenmasse
$k_B = 1.38 \times 10^{-23}$	JK^{-1}	Boltzmann-Konstante
Φ	eV	Barrierenhöhe zwischen Si und SiO_2
A	AV^{-2}	Fowler-Nordheim Konstante
B	mV^{-1}	Fowler-Nordheim Konstante
λ_f	m	Mittlere freie Weglänge eines Elektrons in Silizium
I_{FN}	A	Fowler-Nordheim Strom
I_{HI}	A	Strom der Hot-Elektron-Injection
j_{FN}	$Ac m^{-2}$	Fowler-Nordheim Stromdichte
I_R	A	Lesestrom
I_A	A	Strom im Arbeitspunkt (V_A, I_A)
I_s	A	Stromgröße beim Stresstest
V_A	A	Spannung im Arbeitspunkt (V_A, I_A)
x_T	m	Tunnelstrecke
x_{fW}	m	Mittlere freie Weglänge
x_{DF}	m	Defektfreier Bereich
T	K	Temperatur
ϵ_a	eV	Aktivierungsenergie
t_L	s	Lagerzeit
T_s	s	Stressdauer
t_R	s	Zeitdauer für die Schwellspannungsänderung Δv_{thR}
V_P	V	Programmierspannung
V_E	V	Löschspannung
$V_{P/E}$	V	$V_P = V_E = V_{P/E}$
V_R	V	Lesespannung
V_{CG}	V	Control-Gate Spannung
V_D	V	Drain-Spannung
V_S	V	Source-Spannung
V_G	V	Gate-Spannung
V_B	V	Substrat-Spannung

V_{FG}	V	Floating-Gate Potenzial
V_{tox}	V	Spannung über dem Tunneloxid
v_{th}	V	Schwellspannung
v_{thH}	V	Schwellspannung des programmierten Zustandes
v_{thL}	V	Schwellspannung des gelöschten Zustandes
v_{thW}	V	Schwellspannungsfenster
v_{thW0}	V	v_{thW} ohne Oxidladung
v_{thH0}	V	v_{thH} ohne Oxidladung
v_{thL0}	V	v_{thL} ohne Oxidladung
v_{thN}	V	neutrale Schwelle des Auslesetransistors
S_{thH}	V	Standardabweichung von v_{thH}
S_{thL}	V	Standardabweichung von v_{thL}
$v_{th,max}$	V	Maximale Schwellspannung des Arrays
$v_{th,min}$	V	Minimale Schwellspannung des Array
ΔV_{SG}	V	Spannungsabfall über dem Select-Transistors
ΔV_{dpl}	V	Spannungsabfall über Verarmungszone
A_{inj}	m^2	Injektorfläche
C_{CG}	F	Control-Gate-Kapazität
C_{inj}	F	Kapazität des Injektors zum Substrat
C_B	F	Kapazität zwischen Floating-Gate und Bulk
C_P	F	Parasitäre Kapazitäten
C_T	F	Gesamtkapazität der Speicherzelle
C_{AT}	F	Kapazität des Auslesetransistors
k_D	1	Drain Kopplungsfaktor
k_{CG}	1	Control-Gate Kopplungsfaktor
t_{ox}	m	Tunneloxididicke
x	m	Ladungsschwerpunkt
x_+	m	Ladungsschwerpunkt der negativen Oxidladung
x_-	m	Ladungsschwerpunkt der positiven Oxidladung
x_R	1	Relativer Ladungsschwerpunkt x/t_{ox}
Q_{FG}	C	Floating-Gate Ladung
Q_{ox}	C	Oxidladung
ΔQ_{ox}	Cm^{-2}	Änderung der Oxidladungsdichte
ΔQ_{SS}	Cm^{-2}	Oberflächenladungsdichte
N_{it}	cm^{-2}	Defektdichte
N_{SS}	cm^{-2}	Anzahl der Grenzflächendefekte

N_{gt}	cm^{-2}	Durch Stress erzeugte Defekte
Q_{BD}	Ccm^{-2}	Charge to Breakdown
Q_{inj}	Ccm^{-2}	Akkumulierte injizierte Ladung
σ_i	cm^2	Wirkungsquerschnitt für Defekte
σ_g	cm^2	Wirkungsquerschnitt für neu generierte Defekte
g	s^{-1}	Defekterzeugungsrate
K	1	Modellparameter des Ladungseinbaus
n	1	Modellparameter des Ladungseinbaus
λ	1	Modellparameter

Abkürzungen

Abkürzung	Erläuterung
NVM	Non Volatile Memory
MOS	Metal Oxide Semiconductor
MIMIS	Metal Isolator Metal Isolator Semiconductor
FAMOS	Floating gate Avalanche MOS
SSIMOS	Shielded Substrate Injection MOS
EPROM	Electrically programmable Read Only Memory
EEPROM	Electrically Erasable Programmable Read Only ROM
FLOTOX	FLOating gate Thin OXide
NAND	Not AND
NOR	Not OR
LOCOS	Local Oxidation of Silicon
FIB	Focused Ion Beam
TEM	Transmission Electron Microscope
EDMR	Electrically Detected Magnetic Resonance
WKB	Wentzel Kramers Brillouin
CCS	Constant Current Stress
CVS	Constant Voltage Stress
PVS	Pulsed Voltage Stress
RVS	Ramped Voltage Stress
ERVS	Exponentially Ramped Voltage Stress
SMU	Source Measurement Unit
AFU	Analog Feedback Unit
WLR	Wafer Level Reliability
TAT	Trap Assisted Tunneling

SILC
TDDB

Stress Induced Leakage Current
Time Dependent Dielectric Breakdown

Kurzfassung

In der Automobilindustrie sind nicht flüchtige Speicher NVM (Non volatile Memory) ein wichtiger Bestandteil der Automobilelektronik und werden zur dauerhaften Abspeicherung von Kalibrierdaten und Software benötigt. Hieraus resultieren hohe Anforderungen an den Datenerhalt und die Zykelfestigkeit beim Programmieren und Löschen. Beide Eigenschaften hängen eng mit der Oxidqualität zusammen und können bisher nur in zeitintensiven Messungen bestimmt werden, sodass auf Wafer-Ebene nur exemplarische Messungen praktikabel sind.

Das Ziel dieser Arbeit ist die Entwicklung eines zeitraffenden Testverfahrens zur Bestimmung der Zykelfestigkeit und der Oxidqualität von EEPROM, um Testzeit und Testkosten signifikant zu senken. Da die Zykelfestigkeit von Speicherzellen durch Defekte und Oxidladungen im Siliziumdioxid limitiert ist, basiert das entwickelte Testkonzept auf der Messung und Modellierung des Ladungseinbaues. Durch Untersuchungen zum Ladungseinbau konnte gezeigt werden, dass die Oxiddegradation von Speicherzellen und MOS-Kapazitäten über eine lineare Korrelation zusammenhängen. Hierzu wurde das Tunneloxid der Speicherzelle unter gepulstem (dynamischem) Spannungsstress degradiert, wohingegen die Oxide der MOS-Kapazitäten konstantem (statischem) Spannungsstress ausgesetzt wurden. Durch eine geeignete Modellierung ließ sich die gefundene lineare Korrelation begründen. Dieser Zusammenhang ermöglicht erstmals eine indirekt Bestimmung des Ladungseinbaues von EEPROM über statisch gestresste MOS-Kapazitäten. Auf der Grundlage dieser Ergebnisse konnte ein Testverfahren entwickelt werden, mit dem EEPROM-Eigenschaften bis zu 100-mal schneller evaluiert werden können als über eine direkte Messung an Speicherzellen. Die Testergebnisse ermöglichen eine genaue Beurteilung der Datenwechselstabilität und weichen weniger als 5 % von den direkten Messungen am Bauteil ab. Über die extrahierten Modellparameter lässt sich die Oxiddegradation und der Degradationsmechanismus in den Speicherzellen bewerten. Die Entwicklung des Testverfahrens umfasst folgende Positionen:

- (1) Die Modellierung der Programmier- und Löschdynamik in Abhängigkeit von Layout- und Prozessparametern sowie elektrischen Parametern.
- (2) Die Untersuchung des Ladungseinbaues mit verschiedenen Methoden, um die Oxidalterung in Speicherzellen und Teststruktur zu modellieren und zu vergleichen. Hierbei steht der funktionale Zusammenhang zwischen Oxidladung und Ladungsdurchfluss im Mittelpunkt der Untersuchungen.
- (3) Die Entwicklung eines rechnerischen Verfahrens zur Simulation der Datenwechselstabilität mithilfe von schnellen Stressmessungen.

Abstract

Beside the use of non-volatile memories (NVM) for storing software data permanently they are important components in automotive electronic systems. In these applications, NVM have to fulfill high requirements concerning data retention and program/ erase endurance. Both characteristics, the data retention and the program /erase endurance are correlated with the oxide quality and have to be evaluated in time consuming test procedures. Therefore there is a demand, that the sample size for testing on wafer level, must be kept at low level.

The intention of this PhD-thesis is to develop an accelerated testing procedure for wafer level applications, to forecast the endurance performance of EEPROM. As the endurance performance of EEPROM is limited by the generation of defects and oxide charges in SiO_2 , the new testing procedures based on the measurement of oxide charges in dependence on the injected charge. A detailed investigation on the charge build up in the tunnel oxide of EEPROM as well as on simplified MOS capacitances reveals a linear relation between the oxide charging of both devices. To show this relation, the tunnel oxide of EEPROM is stressed under pulsed voltage stress. Whereas the test capacitances are subjected to constant voltage stress. The charge build up is measured under different conditions and is modelled with established oxide models for charge trapping in order to show the linear correlation between EEPROM and test capacitances. Based on these findings, an accelerated test procedure is developed for endurance lifetime prediction. The procedure based on fast stress tests on MOS-capacitances and affords a more than 100 times faster evaluation of the endurance and oxide performance as directly measured on EEPROM devices. The test results differ less than 5% from measurement results directly measured on 384Bit EEPROM-array structures. In addition to this, the oxide quality can be reviewed by a trapping-yield dependend model parameter. The development of the testing procedures includes the following steps:

- 1) The development of a model for simulating the program / erase behaviour of an EEPROM device in dependence on layout- and process parameters.
- 2) The measurement of the charge build up in the tunneloxide of EEPROM and MOS-capacitances with different methods and under different stress conditions. The oxide charge is measured and modelled in dependence on the injected charge.
- 3) The development of a simulation procedure for lifetime prediction of EEPROM, using stress test results from MOS-capacitance as input data.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Einführung	1
1.2	Einordnung des Testverfahrens	3
2	Grundlagen zu Speicherzellen	5
2.1	Floating-Gate-Speicherzellen	5
2.1.1	Geschichte der Floating-Gate-Speicher	5
2.1.2	Programmier- und Löschmechanismen	8
2.1.3	Aufbau von Single-Poly-Speicherzellen	9
2.2	Programmier- und Löschdynamik von Speicherzellen	12
2.3	Modellentwicklung	14
2.4	Leitungsmechanismen in Isolatoren	16
2.4.1	Theorie der Fowler-Nordheim-Leitung	17
2.4.2	Temperaturabhängigkeit	19
2.4.3	Fowler-Nordheim-Leitung mit Oxidladungen	20
2.4.4	Modell mit Ladungseinbau	21
2.5	Defekte in Siliziumdioxid	23
2.5.1	Defekt-Klassifizierung	23
2.5.2	Isolatorschädigungsmodell	25
2.6	Oxidladungen	28
2.6.1	Ladungseinbau: Kinetische Modelle	28
2.6.2	Ladungseinbau: Potenzgesetzmäßigkeit	30
2.7	Zuverlässigkeitsverhalten von Speicherzellen	32
2.7.1	Datenwechselstabilität	32
2.7.2	Datenerhalt	34
3	Modellierung und Messungen	37
3.1	Transiente Simulation	37
3.1.1	Simulation mit Oxidladungen	41
3.1.2	Modelloptimierung	43
3.2	Messungen	46
3.2.1	Aufbau der Messplätze	46
3.2.2	EEPROM-Teststruktur	48
3.2.3	Programmier- und Löschverhalten	49
3.2.4	Messungen zur Datenwechselstabilität	51

4	Untersuchungen zur Oxiddegradation	55
4.1	Bauelemente der Single-Poly-Speicherzellen	55
4.1.1	Auslesetransistor	56
4.1.2	Select-Transistor	57
4.1.3	Injektor	58
4.2	DiMaria-Methode zur Bestimmung der Oxidladung	61
4.2.1	Stressmessungen	63
4.2.2	Einfluss von Oxiddicke und Temperatur	64
4.3	Ladungseinbau in Speicherzellen (Papadas-Methode)	66
4.4	Diskussion des Ladungseinbaus	69
4.5	Untersuchungen zu Modellparametern	73
4.5.1	Richtungsabhängigkeit	73
4.5.2	Stromdichte- und Temperaturabhängigkeit	73
4.5.3	Korrelation zwischen Speicherzelle und Teststruktur	76
4.6	Entwicklung eines schnellen EEPROM-Testverfahrens	79
4.6.1	Anforderungsprofil an beschleunigte Testverfahren	79
4.6.2	Evaluierung des Ausfallverhaltens über Indikatorgrößen	79
4.6.3	WLR-Test zur Bestimmung der EEPROM-Datenwechselstabilität	81
4.6.4	WLR-Simulationsverfahren	83
4.6.5	Simulationsergebnisse	84
4.6.6	Modelloptimierung	86
5	Zusammenfassung und Diskussion	91
	Literatur	93

1 Einleitung

1.1 Einführung

Seit den 70er-Jahren ist Elektronik ein fester Bestandteil von Automobilen. Durch Innovationen und die stetige Weiterentwicklung in diesem Bereich konnten in den letzten Jahrzehnten die Fahrzeugsicherheit, der Fahrkomfort und die Effizienz erhöht werden. Mittlerweile sind nahezu 90 % aller Innovationen elektronikgetrieben [1], was sich z. B. in der stetig wachsenden Systemkomplexität der Kfz-Elektronik zeigt. Die Zunahme der Systemkomponenten und deren Vernetzung führt zu einem rasanten, exponentiellen Wachstum der Systemkomplexität, siehe Bild 1.1. Diese Entwicklung zeigt, dass mechanische Anwendungen in zunehmendem Maße durch mechatronische und elektronische Applikationen ersetzt werden. Die elektronischen Anwendungen müssen unter extremen Umweltbedingungen operieren, wie z. B. hohen Temperaturen, Luftfeuchtigkeit und mechanischen Vibrationen, sodass besonders hohe Anforderungen an die Robustheit, Zuverlässigkeit und die Lebensdauer der Bauteile gestellt werden. Ein Grund für die strikten Anforderungen der Automobilelektronik, wie sie z. B. im JEDEC-Standard oder im AEC-Q100-Standard zu finden sind, ist die Tatsache, dass mittlerweile 50 % aller Fahrzeugrückrufaktionen durch Elektronikausfälle verursacht werden. Ein weiterer Anstieg auf über 60 % wird für das Jahr 2013 prognostiziert [2]. Im Gegensatz zur

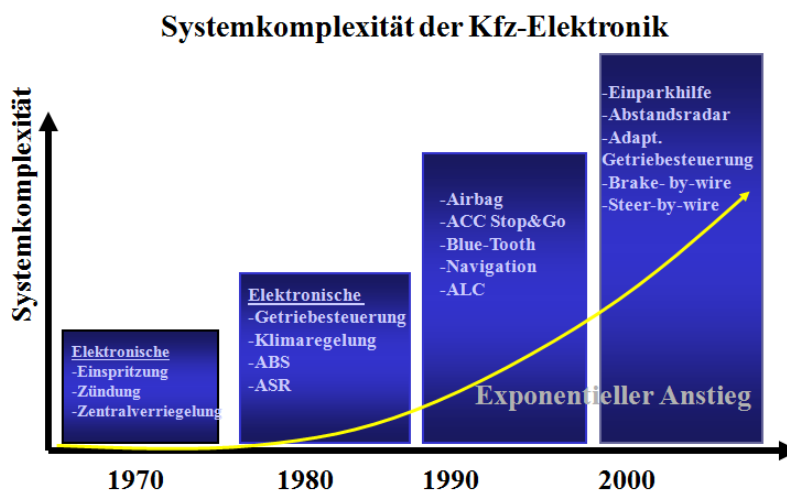


Bild 1.1: Die stetige Zunahme der Automobilelektronik führt zu einem rasantem Wachstum der Systemkomplexität.

„Consumerelectronic“ ist die Automobilelektronik nicht nur durch eine deutlich höhere Lebensdauer charakterisiert, sondern muss auch über einen längeren Zeitraum verfügbar sein als Elektronik aus dem Unterhaltungs- und Telekommunikationsbereich. Der

1 Einleitung

für die "Consumerelectronic" charakteristische Trend zur Miniaturisierung findet in abgeschwächter Form auch in der Automobilelektronik statt, sodass die Packungsdichten der Schaltungen und die Rechenleistung der Prozessoren stetig steigen. Ein hieraus resultierendes Problem ist die Zunahme der Umgebungstemperatur durch die steigende Wärmedissipation der Elektronik, wodurch die Anforderungen an Hochtemperaturanwendungen ($T \geq 125^\circ\text{C}$) [3,4], wie sie z. B. bei der Motor- und Getriebesteuerung zum Einsatz kommen, deutlich steigen. Weitere Bereiche für zukünftige Hochtemperaturanwendung könnten

- die Abgasaufbereitung
- die Bremssysteme
- und die Motorperipherie

sein, was den Bedarf an hochtemperaturfester Elektronik weiter wachsen ließe. Diese hohen Anforderungen an die Zuverlässigkeit von Bauteilen zeigen sich auch bei den nicht flüchtigen Speichern, die bei der dauerhaften Speicherung von Software sowie Kalibrierdaten zum Einsatz kommen. Die Speicherzellen müssen eine Datenspeicherung (Datenerhalt) von 10 bis 20 Jahren garantieren und sich durch eine hohe Zykelfestigkeit (Datenwechselstabilität) beim Programmieren und Löschen auszeichnen. Dies heißt im konkreten Fall, dass Speicherzellen ohne Auftreten eines Funktionsfehlers bis zu 100.000-mal programmierbar und löschar sein müssen. Um diese Spezifikationen erreichen zu können, müssen höchste Ansprüche an die verwendeten Oxide gestellt werden, da der Datenerhalt und die Datenwechselstabilität maßgeblich hiervon abhängen. Daher erfolgen bereits auf Wafer-Ebene Tests zur Bestimmung der Oxidqualität. Hierbei erweisen sich Testverfahren zur Bestimmung der Zykelfestigkeit und des Datenerhalts als äußerst zeitintensiv, weshalb in der Entwicklungs- und Produktionsphase meist nur exemplarische Messungen erfolgen. Die Untersuchung eines kompletten Wafers oder einer Charge (25 Wafer) ist zu zeit- und kostenintensiv. Um dieses Problem zu lösen, wird in dieser Arbeit ein Testverfahren zur Bestimmung der Zykelfestigkeit und Oxidqualität entwickelt, mit dem Testzeit und Testkosten reduziert werden können. Die Arbeit ist folgendermaßen gegliedert:

In **Kapitel 2** werden die Grundlagen zu NVM-Bauelementen und insbesondere der Single-Poly-Speicherzelle bereitgestellt, die im Fokus der Untersuchungen steht. Weiterhin werden Leitungsmechanismen in Isolatoren, Defektypen in SiO_2 und Modelle zur mathematischen Beschreibung des Ladungseinbaus vorgestellt.

In **Kapitel 3** wird ein Model zur transienten Simulation der Programmier- und Löschdynamik von Single-Poly-Speicherzellen entwickelt. Darüber hinaus wird der Einfluss von Oxidladungen auf die Programmierung und Löschung untersucht.

In **Kapitel 4** werden die NVM-relevanten Bauelemente diskutiert, Messmethoden zur Bestimmung von Oxidladungen bereitgestellt und unter verschiedenen Stressbedingungen verifiziert. Die Ergebnisse zum Ladungseinbau ermöglichen die Entwicklung eines neuen Testverfahrens zur Bestimmung der Datenwechselstabilität. Das Testkonzept, die Grundgleichung des Tests und die Simulationsergebnisse werden vorgestellt und diskutiert.

In **Kapitel 5** werden die Ergebnisse der Arbeit zusammenfassend diskutiert und es wird ein Ausblick auf weiterführende Untersuchungen gegeben.

1.2 Einordnung des Testverfahrens

Das hier entwickelte Testverfahren für EEPROM wird auf Wafer-Ebene eingesetzt und ermöglicht die Messung der Datenwechselstabilität und der Oxidqualität in EEPROM Bauteilen. Das Testverfahren basiert auf Stressmessungen, die an Tunneloxidkapazitäten erfolgen, die durch Anlegen einer konstanten Spannung beschleunigt degradieren. Die zentralen Messgrößen sind hierbei nicht die klassischen TDDB-Werte, wie die Durchbruchladung oder die Zeitdauer bis zum Oxiddurchbruch, sondern die Oxidladungen als Funktion des Ladungsdurchflusses. Mithilfe eines neuen rechnerischen Verfahrens können hierüber EEPROM-Eigenschaften über 100-mal schneller bestimmt werden als über eine direkte Messung an EEPROM-Bauteilen.

Für eine bessere Einordnung des Testverfahrens soll ein kurzer Überblick über Testverfahren und die dazugehörigen Komponenten der elektronischen Baugruppen gegeben werden. Die Zuverlässigkeit von elektronischen Baugruppen im Automobilbereich umfasst das Gehäuse, die Verbindungstechnik und die Bauelemente. Wie aus Bild 1.2 hervorgeht, finden die Zuverlässigkeitstests für einzelne Bauelemente bereits auf Wafer-Ebene statt und erfassen z. B. die Oxidqualität (Hot-Carrier-Injektion, TDDB-Messungen) und die Leiterbahnen (Elektromigration). Nach der Integration der Bauele-

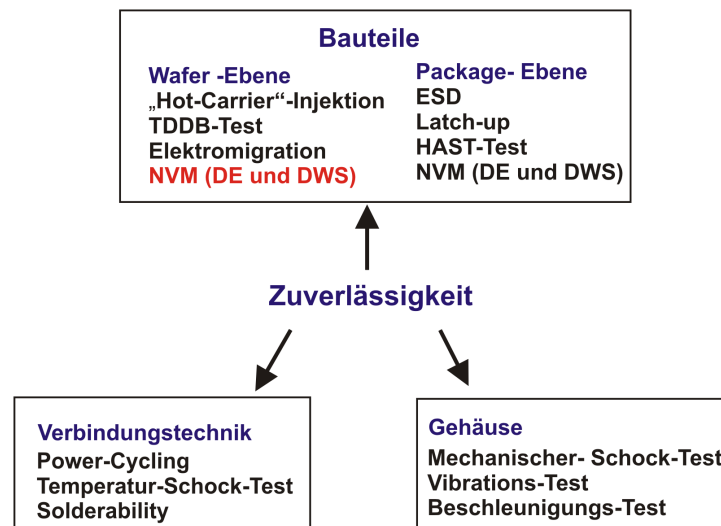


Bild 1.2: Die Zuverlässigkeit von elektronischen Baugruppen umfasst die Verbindungstechnik, das Gehäuse und die Bauteile bzw. Schaltungen. Die hier aufgeführten Testverfahren sind im Teststandard AEC-Q100-REV-G beschrieben.

mente in ein Chipgehäuse oder Modul (Package-Ebene) finden weitere Tests statt. Hierin werden das ESD (Electrostatic Discharge) und das Latch-up-Verhalten der Schaltungen untersucht. Die nicht flüchtigen Speicher NVM werden auf Package-Ebene gesondert untersucht, da hierbei die bauelementespezifischen Eigenschaften wie der Datenerhalt (DE) und die Datenwechselstabilität (DWS) von Interesse sind. Um die Zuverlässigkeit der Verbindungstechnik (Bonddrähte, Chip- und Systemlötlung) zu untersuchen, gibt es im AEC-Q100-REV-G-Teststandard für „automotive“ Elektronik zahlreiche Tests, wie z. B. den „Power-Cycling“-Test oder den Temperatur-Schock-Test

1 Einleitung

(TST). Aufgrund von thermischer und mechanischer Wechselbelastung degradieren sowohl Bondverbindung und Lotschicht mit zunehmender Stressdauer und -belastung. Insbesondere bei Lotschichten bilden sich Risse und Hohlräume aus, die zu einer verminderten Wärmeleitfähigkeit bzw. Überhitzung des Bauelementes führen. Weitere Untersuchungen betreffen die Zuverlässigkeit des Gehäuses und umfassen Tests zum Verhalten unter Vibration, Beschleunigung oder mechanischem Schock.

2 Grundlagen zu Speicherzellen

2.1 Floating-Gate-Speicherzellen

2.1.1 Geschichte der Floating-Gate-Speicher

Der erste nicht flüchtige Speicher wurde 1967 von D. Kahng und S. M. Sze [5] vorgestellt. Diese Art von Speicher basiert auf Halbleiterbauelementen und ermöglicht die dauerhafte Speicherung von Informationen ohne Anlegen von externen Spannungen. Grundlage hierfür war eine MOS-Struktur, bei der das Gate durch eine Schichtfolge aus Siliziumdioxid I_1 , Metall M_1 , Siliziumdioxid I_2 und Metall M_2 (Control-Gate) ersetzt wurde, Bild 2.1a). Hieraus leitet sich auch die Abkürzung MIMIS (Metal Insulator Metal Insulator Semiconductor) ab. Da die leitfähige Schicht M_1 keinen externen Kontakt besitzt und durch die Isolatoren I_1 und I_2 isoliert ist, wird sie auch als Floating-Gate (float engl. schweben) bezeichnet. Durch positive oder negative Aufladung des Floating-Gates können die Transferkennlinien bzw. die Schwellspannungen des MOS-Transistors verschoben werden, sodass zwischen zwei unterschiedlichen Zellzuständen „0“ und „1“ unterschieden werden kann. Die Schwellspannung im programmierten (gelöschten) Zustand wird mit v_{thH} und (v_{thL}) bezeichnet.

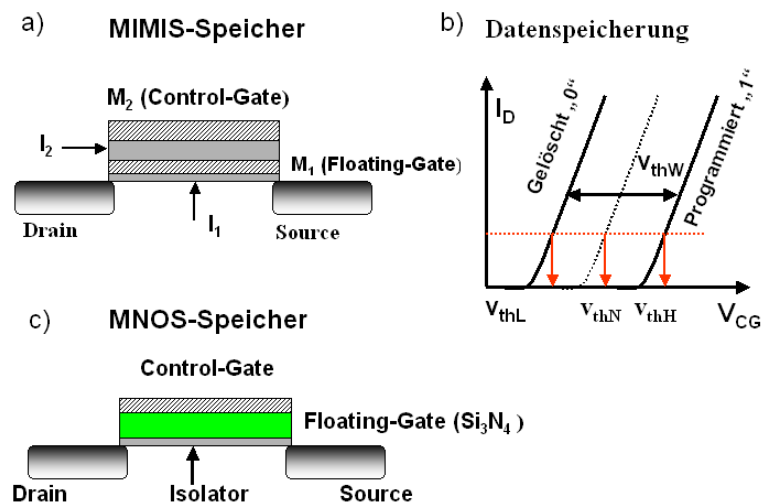


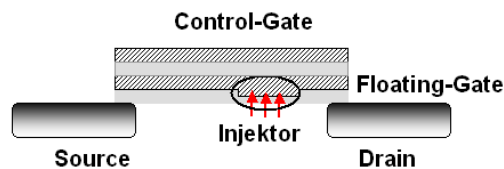
Bild 2.1: a) Schematische Darstellung der ersten Floating-Gate-Speicherzelle b) Schematische Darstellung der Transferkennlinien für unterschiedliche Zellzustände c) Zeigt den schematischen Aufbau des ersten nitridbasierten Speichers.

2 Grundlagen zu Speicherzellen

Die Schwellspannung im neutralen Ladungszustand des Floating-Gates ist v_{thN} . Der Abstand zwischen den Transferkennlinien wird als Schwellspannungsfenster v_{thW} bezeichnet, Bild 2.1b).

Der MIMIS-Speicher wurde elektrisch programmiert und gelöscht, was in beiden Fällen über direktes Tunneln geschah. Hierfür musste die Isolatorschicht I_1 dünner als 5 nm sein, um direktes Tunneln anwenden zu können. Da Ende der 60er-Jahre die Qualität von Oxiden ≤ 10 nm noch gering war, erreichten die MIMIS-Zellen einen mäßigen Datenerhalt, da sich das Floating-Gate schnell über Schwachstellen im Oxid entladen konnte [6]. Daher ging von der ersten Speicherzelle ein eher geringer praktischer Nutzen aus. Die historische Bedeutung ist allerdings als hoch einzustufen, da diese erste Speicherzelle alle Merkmale heutiger Floating-Gate-Speicher aufwies. In der Folgezeit wurde das MIMIS-Speicherkonzept weiter verbessert und modifiziert, woraus die Floating-Gate-Speicher und die nitridbasierten Speicher hervorgingen, Bild 2.1c). So wurde z. B. die Oxiddicke erhöht und im Fall der nitridbasierten Speicher das leitfähige Floating-Gate-Material durch Silizium-Nitrid ersetzt, was zu den MNOS-Speichern (Metal Nitride Oxide Semiconductor) führte, die erstmals von Wegner et al. [7] 1967 vorgestellt wurden. Bei diesen Speicherzellen wird im Programmier- und Löschfall am Control-Gate eine positive bzw. negative Spannung angelegt, um Ladungsträger in die Si_3N_4 -Schicht zu injizieren und dort grenzflächennah zu speichern. Die ersten Floating-Gate-Speicher von praktischer Bedeutung waren die von Frohman-Bentchkowsky entwickelten FAMOS (Floating-Gate Avalanche MOS) [8,9], die zur Kategorie der EPROM-Speicher (Electrically Programmable Read Only Memory) gehören. Die EPROM können elektrisch durch Fowler-Nordheim-Injektion [10,11,14]

a) UV-Löschbares EPROM



b) FLOTOX- EEPROM

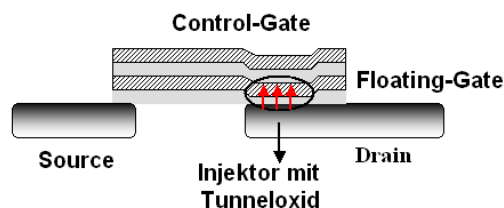


Bild 2.2: a) Schematische Darstellung eines Querschnitts durch eine UV-löschbare EPROM-Speicherzelle. Die Programmierung erfolgt über Fowler-Nordheim-Injektion oder mithilfe von „heißen Elektronen“. b) Querschnitt durch eine FLOTOX-Speicherzelle.

oder über heiße Elektronen [12,13,14] programmiert werden, wobei das Löschen über UV- oder Röntgenstrahlung geschieht und bis zu 20 Minuten in Anspruch nehmen kann

[6]. In Bild 2.2a) ist eine EPROM-Speicherzelle mit Trench-Gate-Oxide-Struktur gezeigt. Floating-Gate und Control-Gate sind übereinander angeordnet und bestehen aus hoch dotiertem Polysilizium. Eine Speicherzelle mit zwei übereinander angeordneten Polysiliziumschichten wird im folgenden auch als Doppel-Poly-Speicherzelle bezeichnet. Deutlich ist der Bereich mit herabgesetzter Oxiddicke zu erkennen, über den die Programmierung stattfindet und der im Folgenden als Injektor bezeichnet wird. Die Löschung erfolgt über die gesamte Kanalbreite des Bauteils.

Die hohe Zeitdauer für die Löschung der Speicherzellen war ein Grund dafür, dass sich aus den EPROM die EEPROM (**E**-Erasable-**P**ROM) entwickelt haben, die sich elektrisch programmieren und löschen lassen. Eine bekannte EEPROM-Speicherzelle ist die von Harari et al. [15] entwickelte FLOTOX-Speicherzelle (**F**loating gate **T**hin **O**Xide), die schematisch in Bild 2.2b) dargestellt ist. Bei der FLOTOX-Speicherzelle befindet sich der Injektor über dem Draingebiet und weist eine Oxiddicke von ca. 10nm auf (Tunneloxid). Das Programmieren sowie das Löschen der Speicherzelle basiert auf der Fowler-Nordheim-Injektion. Beim Programmieren wird am Control-Gate eine Programmiervoltage V_P (**P**rogramm) angelegt; Drain und Bulk liegen dabei auf Ground und der Source-Kontakt ist floatend. Durch die kapazitive Kopplung zwischen Control-Gate und Floating-Gate baut sich über dem Tunneloxid ein elektrisches Feld auf, das eine Fowler-Nordheim-Injektion bzw. negative Aufladung des Floating-Gates bewirkt. Im Löschfall wird am Drainkontakt eine Löschspannung V_E (**E**rase) angelegt; Control-Gate und Bulk liegen auf Ground und der Source-Kontakt ist floatend. Der Fowler-Nordheim-Tunnelstrom fließt vom Floating-Gate ins Draingebiet und lädt das Floating-Gate positiv auf. Nahezu der gesamte Tunnelstrom fließt beim Programmieren und Löschen durch das Tunneloxid des Injektors.

Weitere EEPROM-Speicherzellen sind die SSIMOS (**S**hielded **S**ubstrate **I**njection **M**OS) und die Single-Poly-Speicherzelle, bei denen möglichst hohe Kopplungsverhältnisse zwischen Control-Gate und Floating-Gate angestrebt werden [6]. Bei der SSIMOS Zelle ist das Floating-Gate nahezu vollständig vom Control-Gate umschlossen, um die Kapazität zwischen Control-Gate und Floating-Gate zu maximieren. Im Fall der Single-Poly Speicherzellen wird die Kapazität zwischen Control-Gate und Floating-Gate von einer großflächigen MOS-Kapazität mit Tunneloxid geliefert.

Aus den EPROM bzw. EEPROM-Speicherzellen gingen die Flash-Speicherzellen hervor, die 1984 von Masuoka et al. auf dem „International Electron Device Meeting“ präsentiert wurden [16,17]. Die Flash-Speicherzellen sind EEPROM-Speicher und daher elektrisch programmierbar und löschar. Eine Besonderheit gegenüber EEPROM betrifft die Anordnung der Zellen in Arrays. Hierin werden die Zellen zu Blöcken zusammengefasst, in denen sämtliche Flash-Speicherzellen gleichzeitig gelöscht werden können. Für die Anordnung von Flash-Speicherzellen innerhalb eines Arrays wurde 1985 von Masuoka eine NAND-Flash-Zelle vorgeschlagen [18], durch die eine deutliche Reduktion der Speichergröße erreicht werden konnte, ohne hierzu die Zellgröße verkleinern zu müssen. Dies war die Grundlage für die heutigen Flash-Applikationen in USB-Sticks, Festplatten, digitalen Fotos und MP3. Die NAND-Flash-Speicher ermöglichen einerseits eine hohe Speicherdichte und andererseits schnelles Löschen und Programmieren von sequenziell abgespeicherten Daten. Weitere Schaltungskonzepte für Flash-EEPROM sind z. B. das AND, DINOR [6] und NOR [6]. Die NOR-Flash-Speicher besitzen gegenüber den NAND-Flash-Speichern den Vorteil einer hohen Aus-

2 Grundlagen zu Speicherzellen

lesegeschwindigkeit und werden primär in Applikationen eingesetzt, die einen schnellen Datenzugriff erfordern, wie z. B. in Mobiltelefonen, Navigationssystemen und Routern.

2.1.2 Programmier- und Löschmechanismen

Für die Programmierung und Löschung stehen verschiedene Leitungsmechanismen zur Verfügung, wie z. B. die Fowler-Nordheim-Injektion und das direkte Tunneln. Ein weiterer Lademechanismus ist die Ladungsträgerinjektion mittels „heißer Elektronen“. Das **direkte Tunneln** hat für die Programmierung und Löschung heutiger Speicherzellen keine Bedeutung. Wie bereits im Zusammenhang mit der ersten Speicherzelle erwähnt, müssen die Tunneloxide dünner als 5 nm sein, um direktes Tunneln anwenden zu können. Da die heute gebräuchlichen nicht flüchtigen Speicher einen Datenerhalt von mehreren Jahren zu gewährleisten haben, werden Tunneloxide zwischen 8-10 nm verwendet, sodass direktes Tunneln nicht anwendbar ist.

Ein Leitungsmechanismus, der bei heutigen EEPROMs und Flash-EEPROMs Anwendung findet, ist die **Fowler-Nordheim-Injektion**. So findet z. B. bei den FLOTOX-Speichern das Programmieren und Löschen ausschließlich über diesen Leitungsmechanismus statt. Auch die Single-Poly-EEPROM, die detailliert in Kap 2.1.3 vorgestellt werden und Gegenstand der Arbeit sind, lassen sich über Fowler-Nordheim-Injektion programmieren und löschen. Hierzu wird über dem Tunneloxid ein elektrisches Feld von mindestens $E=10\text{MVcm}^{-1}$ erzeugt [6], sodass eine feldunterstützte, quantenmechanische Emission stattfindet, Bild 2.3a). Dieser Leitungsmechanismus wird als Fowler-Nordheim-Injektion bezeichnet und lässt sich durch die Fowler-Nordheim-Gleichung beschreiben.

$$j_{FN} = AE^2 e^{-\frac{B}{E}} \quad (2.1)$$

Hierbei ist j_{FN} die Fowler-Nordheim-Stromdichte und E ist das elektrische Feld über

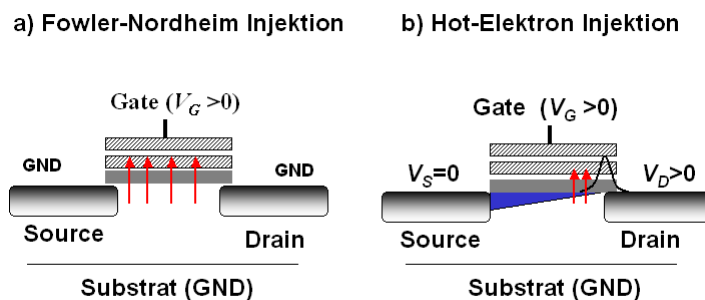


Bild 2.3: a) Fowler-Nordheim-Injektion: Durch Anlegen einer Gatespannung kann eine feldunterstützte, quantenmechanische Emission stattfinden, die das Floating-Gate auflädt. b) Zeigt die Programmierung mittels „heißer“ Elektronen.

dem Oxid. Die Größen A und B werden als Fowler-Nordheim-Konstanten bezeichnet. Die Ladungsinjektion durch "**heiße Elektronen**" (Hot-Elektron-Injektion) ist ein weiterer Leitungsmechanismus, der ausschließlich bei der Programmierung von Speicher-

zellen zum Einsatz kommt. Dem MOS-Transistor der Speicherzelle wird hierzu eine hohe elektrische Feldstärke eingeprägt, wodurch am drainseitigen Ende des n-Kanals Elektronen mit ausreichend hoher Energie entstehen, um ins Floating-Gate zu gelangen, Bild 2.3 b). Der Emmissionsvorgang kann z. B. durch die „Temperature-Theory“ [19] oder das „Lucky Electron“-Modell von C. Hu [20,21] beschrieben werden. Beim „Lucky Electron“-Modell wird angenommen, dass die Elektronen am Kanalende genügend Energie im elektrischen Feld aufnehmen können, um die Potenzialbarriere zwischen Silizium und Siliziumdioxid überwinden zu können. Durch Streuprozesse zwischen den Ladungsträgern und dem Kristallgitter (Phonon-Streuung) werden die Elektronen in Richtung Floating-Gate gestreut. Der mit diesem Modell berechnete Gate-Strom ist direkt proportional zum Drainstrom I_D und hat folgende Form:

$$I_{HI} \propto I_D e^{\frac{\Phi}{2E_m\lambda_f} + \frac{2\Phi}{3E_m\lambda_f}} \quad (2.2)$$

Hierbei steht j_{HI} für den Ladestrom, E_m bezeichnet hier das elektrische Feld am Kanalende und Φ steht für die Potenzialbarriere zwischen Silizium und Gate. Die mittlere freie Weglänge des Elektrons wird λ_f genannt.

2.1.3 Aufbau von Single-Poly-Speicherzellen

Eine Single-Poly-Zelle kann anschaulich aus einer Doppel-Poly-Speicherzelle erzeugt werden, indem die zweite Polysiliziumschicht umgeklappt wird, Bild 2.4. Unterhalb

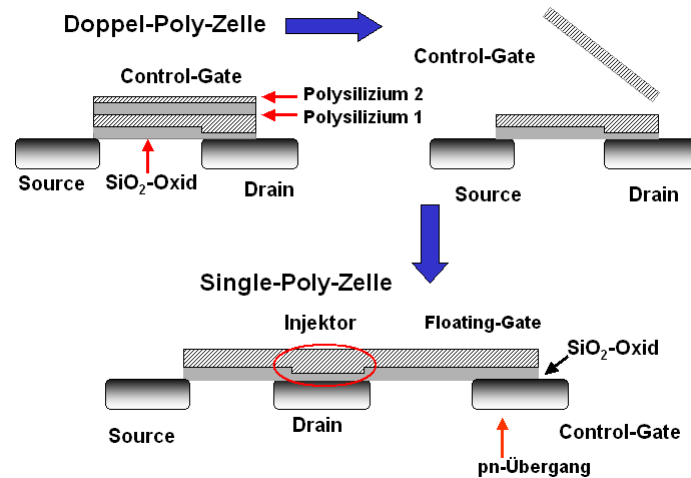


Bild 2.4: Schematischer Übergang von einer Doppel-Poly-Speicherzelle zu einer Single-Poly-Speicherzelle.

der herumgeklappten Polysiliziumschicht befindet sich die n-Wanne des Control-Gates, über die das Floating-Gate kapazitiv gesteuert wird. Das Control-Gate der Single-Poly-Speicherzelle wird gegenüber dem Substrat durch einen pn-Übergang isoliert, Floating-Gate und Control-Gate sind hierbei über Siliziumdioxid voneinander getrennt. Die Programmierung und Löschung findet über das Tunneloxid des Injektors statt und erfolgt analog zur FLOTOX- Speicherzelle. Das schematische Layout der in dieser Arbeit un-

2 Grundlagen zu Speicherzellen

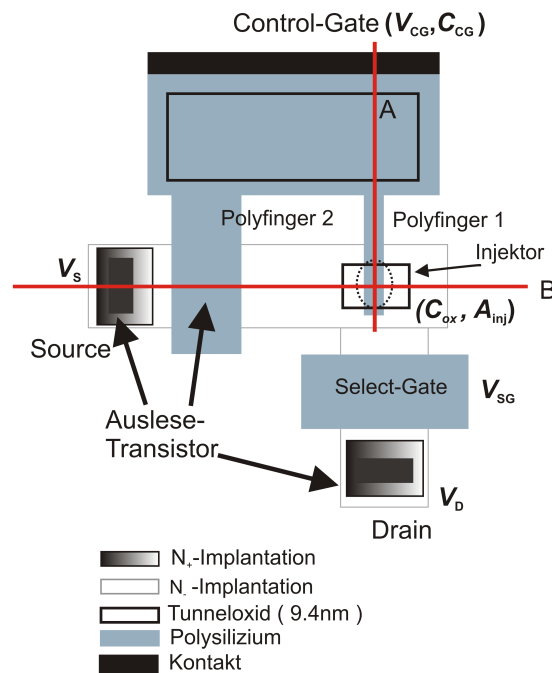


Bild 2.5: Schematischen Darstellung einer Single-Poly Speicherzelle. Die wichtigsten Komponententen der Speicherzelle sind der Auslesetransistor, der Injektor und das Floating-Gate. Der Select-Transistor besteht aus Select-Gate, Source und Drain.

versuchten Single-Poly-Zelle ist in Bild 2.5 skizziert. Die Speicherzellen wurden am Fraunhofer IMS in Duisburg mit einer 350 nm LOCOS-Technologie (Local Oxidation of Silicon) hergestellt. Das hierfür verwendete thermisch gewachsene Tunneloxid hat eine Dicke von 9.4 nm, die Gateoxiddicke beträgt 45 nm und das Feldoxid zur lateralen Isolation besitzt eine Dicke von 240 nm. Die Single-Poly-Speicherzelle ist durch eine Polysiliziumschicht (Floating-Gate) gekennzeichnet und in zwei Bereiche unterteilbar, die hier als Polyfinger 1 und Polyfinger 2 bezeichnet werden. Hiervon getrennt ist der Select-Transistor mit seinem Gate aufgeführt. Der Polyfinger 2 bildet mit Source und Drain den Auslesetransistor, der über das Control-Gate steuerbar ist. Unter dem Polyfinger 1 befindet sich der Injektor, über dessen Tunneloxid die Zelle programmiert und gelöscht wird. Ein weiterer Bereich mit Tunneloxid befindet sich im Bereich des Control-Gates, um eine hinreichend hohe kapazitive Kopplung zwischen Control-Gate und Injektor zu realisieren. Die Kapazitäten zwischen Substrat und Injektor bzw. (Control-Gate) werden als Injektorkapazität C_{inj} und Control-Gate-Kapazität C_{CG} bezeichnet. Die Injektorfläche, durch die die Fowler-Nordheim-Injektion stattfindet, wird A_{inj} genannt.

Die Programmierung und Löschung der Speicherzelle geschieht über Fowler-Nordheim-Injektion. Für die Programmierung der Speicherzelle wird am Select-Gate und am Control-Gate die Programmiervoltage V_P angelegt, wohingegen die Drain- und Bulkspannung auf Ground liegen. Es gelten folgende Bedingungen:

$V_{CG} = V_{SG} = V_P$, $V_B = V_D = \text{GND}$ und $V_S = \text{float}$. Aufgrund der kapazitiven Kopplung

zwischen Control-Gate und Injektor steigt die Spannung über dem Tunneloxid des Injektors an. Die hieraus resultierende Fowler-Nordheim-Injektion lädt das Floating-Gate negativ auf.

Für die Löschung der Zelle wird am Select-Gate und am Drainkontakt die Löschspannung V_E angelegt. Control-Gate und Bulk liegen auf Ground. Zusammenfassend gilt: $V_D = V_{SG} = V_P$, $V_B = V_{CG} = \text{GND}$ und $V_S = \text{float}$.

Über den n-Kanal des Select-Transistors wird die Drainspannung an den Injektor geleitet, wodurch die Spannung über dem Tunneloxid ansteigt und die einsetzende Fowler-Nordheim-Injektion das Floating-Gate positiv auflädt. Da die Single-Poly-Speicherzellen in dieser Arbeit symmetrisch programmiert und gelöscht werden, soll nachfolgend $V_P = V_E = V_{P/E}$ gesetzt werden. Die Spannung $V_{P/E}$ soll als P/E-Spannung bezeichnet werden.

Eine weitere Zelloperation ist neben dem Programmieren und Löschen das Auslesen der Speicherzelle. Hierzu wird $V_S = 0 \text{ V}$, $V_D = 0.5 \text{ V}$ und $V_B = \text{GND}$ gesetzt.

Am Control-Gate wird die Lesespannung V_R angelegt, der korrespondierende Drainstrom gemessen und mit einem Referenzwert I_R verglichen. Für die Zellzustände soll gelten:

$$\begin{aligned} I_D(V_R) < I_R &\Rightarrow \text{„1“} \quad \text{programmierter Zustand} \\ I_D(V_R) > I_R &\Rightarrow \text{„0“} \quad \text{gelöschter Zustand} \end{aligned}$$

Umgekehrt lässt sich zu einem spezifizierten Referenzstrom die Control-Gate-Spannung messen und als Schwellspannung des jeweiligen Zellzustandes definieren.

Zur Visualisierung der Single-Poly-Speicherzelle wurden FIB-Schnitte (Focused-Ion-Beam) entlang der Schnittlinien A und B aus Bild 2.5 durchgeführt. Die TEM-Aufnahmen (Transmission Electron Microscope) der Querschnitte sind in Bild 2.6 und Bild 2.7 zu sehen. Auf Bild 2.6 a) ist der Querschnitt durch Injektor und Polyfinger 2 zu sehen (Schnittgerade B). Die Teilausschnitte b) und c) zeigen vergrößert die Injektorstruktur

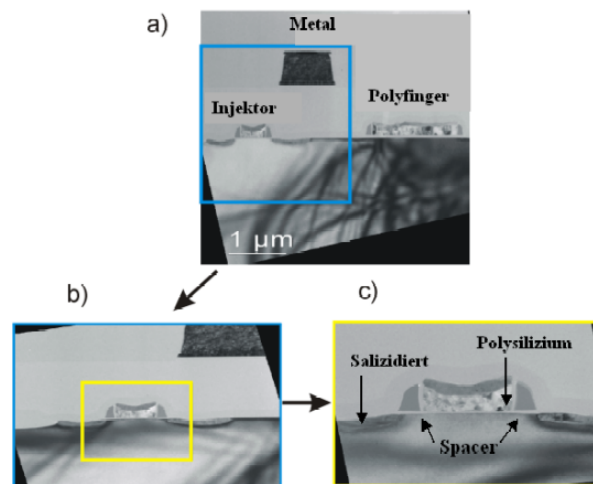


Bild 2.6: a) Zeigt einen TEM-Schnitt entlang der Schnittgeraden B, die den Injektor und den Polyfinger 2 zeigt. In den vergrößerten Ausschnitten b) und c) sind Details des Injektors erkennbar.

2 Grundlagen zu Speicherzellen

mit dem Tunneloxid und der Polysiliziumschicht, die beidseitig durch Spacerstrukturen [22,23] begrenzt wird. Auf dem Polysilizium und neben den Spacern befinden sich salizidierte Bereiche [24], die dunkler erscheinen. Die TEM-Aufnahme in Bild 2.7a) zeigt einen Querschnitt durch das Control-Gate und den Polyfinger 1 (Schnittgerade A). Rechts befindet sich der EEPROM-Injektor und links das Control-Gate; das Polysilizium wird zwischen Injektor und Control-Gate über Feldoxid geleitet. Die Teilbilder b) und c) zeigen den Injektorrand und den Übergang zwischen Tunneloxid und Gateoxid.

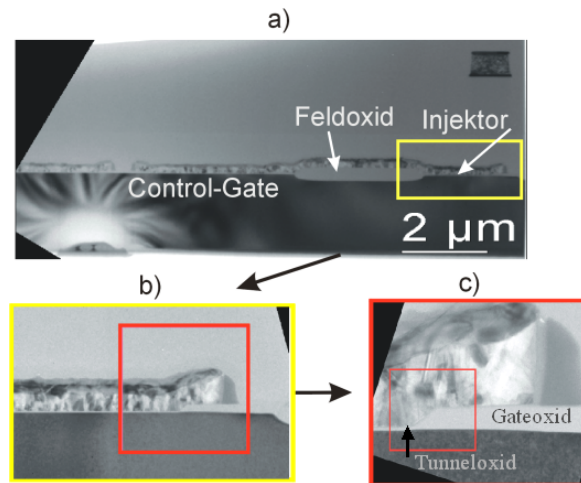


Bild 2.7: Der Querschnitt entlang der Schnittgeraden A zeigt den Injektor und das Control-Gate. a) Der Injektor befindet sich im rechten Teil des Bildes. Links hiervon befindet sich das Control-Gate. Die vergrößerten Ausschnitte b) und c) zeigen das Polysilizium, die Spacer und den Übergang zwischen Gate- und Tunneloxid.

2.2 Programmier- und Löschdynamik von Speicherzellen

Der bereits qualitativ erläuterte Programmier- und Löschvorgang soll in diesem Kapitel modelliert werden. Die Modellierung basiert auf dem Ersatzschaltbild der Speicherzelle (Kapazitätsmodell) und ermöglicht die mathematische Beschreibung der Programmier- und Löschvorgänge.

Programmieren:

Das Bild 2.8 gibt die schematische Darstellung einer Single-Poly-Speicherzelle im Programmierfall wieder. Beim Programmieren wird am Control-Gate und am Select-Gate ein zeitabhängiger Spannungspuls der Höhe V_P angelegt.

Es gilt: $V_{SG} = V_{CG} = V_P$ und $V_B = V_D = \text{GND}$, wobei V_S floatend ist.

Über den kapazitiven Spannungsteiler wird die Control-Gate-Spannung so aufgeteilt, dass über dem Tunneloxid der Anteil V_{ox} abfällt. Das korrespondierende elektrische Feld E_{ox} bewirkt eine Fowler-Nordheim-Injektion, die das Floating-Gate negativ auflädt. Durch den Anstieg der Floating-Gate-Ladung ändert sich die Schwellspannung und das Floating-Gate-Potenzial V_{FG} . Des Weiteren bewirkt die Floating-Gate-Ladung

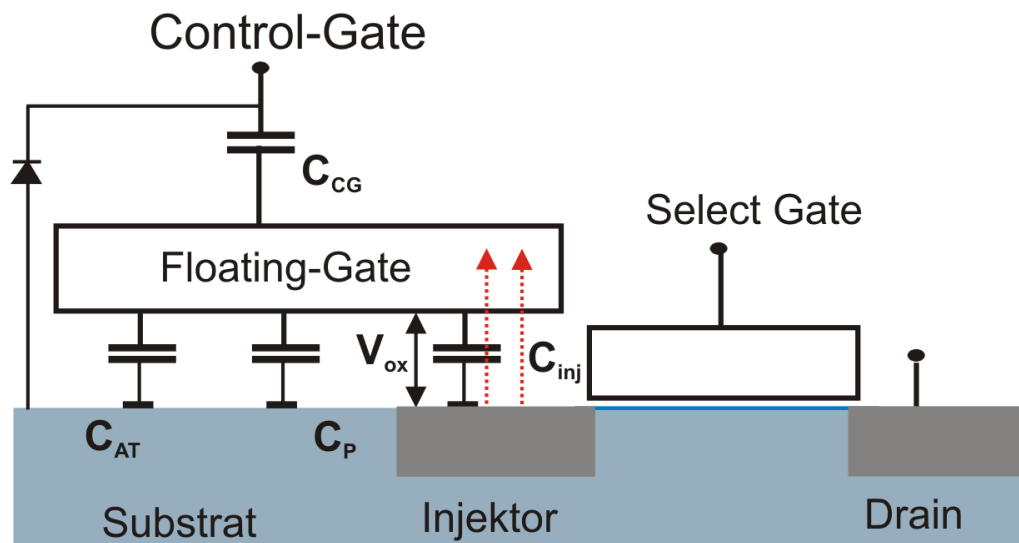


Bild 2.8: Das Ersatzschaltbild der Speicherzelle zeigt den Programmierfall. Die Kapazität C_{AT} bezeichnet die Kapazität des Auslesetransistors, C_{inj} und C_{CG} sind die Injektorkapazität und die Control-Gate-Kapazität. Die parasitären Kapazitäten werden C_P genannt.

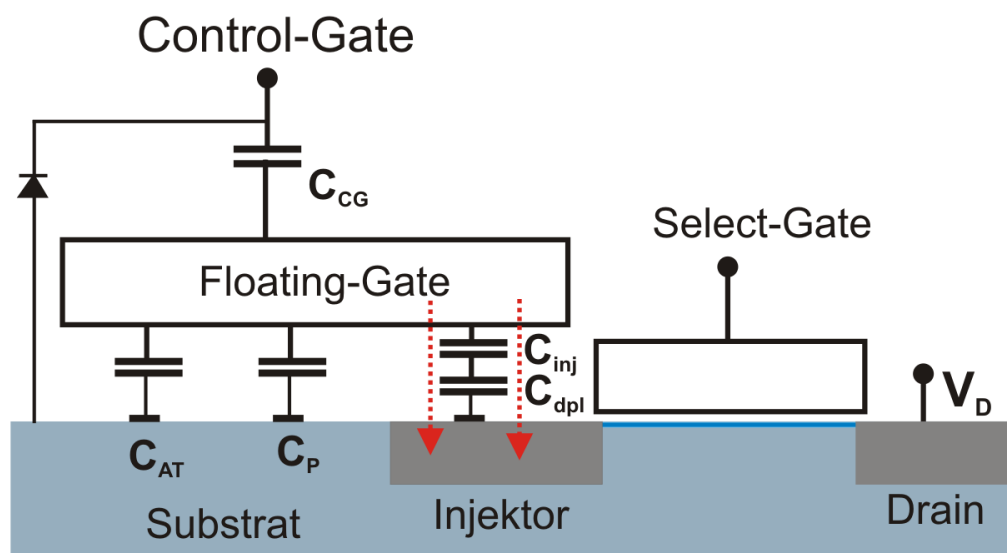


Bild 2.9: Das Ersatzschaltbild der Speicherzelle zeigt den Löschfall. Am Drainkontakt und Select-Gate wird die Löschspannung V_E angelegt. Die Verarmungszone unterhalb des Injektorgebietes wird im Schaltbild durch die Kapazität C_{dpl} berücksichtigt.

2 Grundlagen zu Speicherzellen

den Aufbau eines Gegenfeldes, das die Spannung über dem Tunneloxid V_{ox} verringert und dem Aufladevorgang entgegenwirkt. Die Speicherzelle kann somit über die Floating-Gate-Ladung dynamisch auf externe Felder reagieren.

Löschen

Beim Löschen wird am Select-Gate und Drainkontakt die Spannung V_E angelegt, wobei $V_B = V_{CG} = \text{GND}$ und V_S floatend ist. Der n-Kanal des Select-Transistors leitet die Drainspannung unter den Injektor, sodass die einsetzende FN-Injektion zu einer positiven Aufladung der Speicherzelle führt, Bild 2.9. Im Gegensatz zum Programmieren ist beim Löschen ein zusätzlicher Spannungsabfall über dem Select-Transistor V_{SG} zu beachten [25]. Darüber hinaus kommt es unterhalb des Injektorgebietes zur Ausbildung einer Verarmungszone [25], was in Bild 2.9 durch eine zusätzliche Kapazität C_{dpl} berücksichtigt wird. Der Spannungsabfall über der Verarmungszone wird mit V_{dpl} bezeichnet. Damit ergibt sich für die effektive Löschspannung der folgende Ausdruck:

$$V_E^{eff} = V_D - \Delta V_{SG} - \Delta V_{dpl}$$

Hiermit kann das elektrische Feld über dem Tunneloxid E_{ox} mithilfe des Floating-Gate-Potenzials V_{FG} , der effektiven Drainspannung und der Tunneloxiddicke t_{ox} bestimmt werden

$$E_{ox} = [V_{FG} - V_E^{eff}] / t_{ox}$$

2.3 Modellentwicklung

Ausgangspunkt der Modellierung ist das Ersatzschaltbild (Kapazitätsmodell) in Bild 2.10, auf das der Ladungserhaltungssatz (Gl. 2.3) angewendet wird.

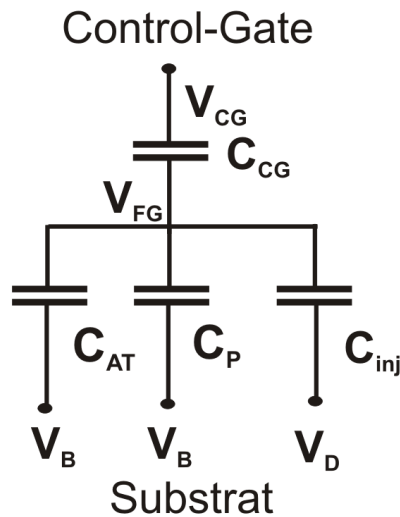


Bild 2.10: Kapazitätsmodell einer Single-Poly-Speicherzelle: Dargestellt ist der Programmierfall.

$$Q_{FG} - \sum_{m=1}^N C_m (V_{FG} - V_m) = 0 \quad (2.3)$$

Hierbei bezeichnet V_m die Spannungen zwischen dem Floating-Gate und seiner Umgebung. Die Kapazitäten zwischen dem Floating-Gate und seinem Umfeld werden durch C_m berücksichtigt. Die Grundgleichungen zur Modellierung von Speicherzellen finden sich z. B. in [26-28] und können über den Ladungserhaltungssatz in Verbindung mit dem Kapazitätsmodell gefunden werden. Die Anwendung von Gl.2.3 auf das Ersatzschaltbild in Bild 2.10 liefert:

$$Q_{FG} + [C_{CG}(V_{FG} - V_{CG}) + C_{AT}(V_{FG} - V_B) + C_P(V_{FG} - V_B) + C_{inj}(V_{FG} - V_{inj})] = 0$$

Hierbei ist C_{AT} die Kapazität des Auslesetransistors, C_{CG} und C_{inj} sind die Control-Gate-Kapazität und die Injektorkapazität. Die parasitären Kapazitäten werden mit C_P bezeichnet. Die Gesamtkapazität heißt C_T ($C_T = C_{inj} + C_{CG} + C_P + C_{AT}$)

Mit diesem Ausdruck können Gleichungen für den Programmier- und Löschfall abgeleitet werden, die das Floating-Gate-Potenzial in Abhängigkeit von der Floating-Gate-Ladung und der externen Spannung beschreiben. Die Größen im Programmier- und Löschfall werden im Weiteren mit einem P bzw. (E) gekennzeichnet.

Programmieren: Wenn $V_{CG} = V_P$, $V_D = GND$ und V_S floatend ist, dann folgt:

$$V_{FG}^P = k_{CG}V_{CG} - \frac{Q_{FG}^P}{C_T} \quad \text{mit} \quad k_{CG} = \frac{C_{CG}}{C_T} \quad (\text{Control - Gate - Kopplungsfaktor}) \quad (2.4)$$

Löschen: Wenn $V_D = V_E$, $V_{CG} = GND$ und V_S floatend, dann folgt:

$$V_{FG}^E = k_D V_D - \frac{Q_{FG}^E}{C_T} \quad \text{mit} \quad k_D = \frac{C_{inj}}{C_T} \quad (\text{Drain - Kopplungsfaktor}) \quad (2.5)$$

Der Control-Gate-Kopplungsfaktor bestimmt im Programmierfall das Verhältnis zwischen der Control-Gate-Spannung und der Spannung über dem Tunneloxid.

$$V_{ox} : V_P = k_{CG} \quad \text{für} \quad Q_{FG} = 0$$

Mit den Gleichungen 2.4 und 2.5 können die elektrischen Felder E_{ox}^E und E_{ox}^P über dem Tunneloxid berechnet werden.

$$E_{ox}^P = \frac{V_{FG}(t)}{t_{ox}} = \frac{V_{CG}(t)k_{CG} - \frac{Q_{FG}^P(t)}{C_T}}{t_{ox}} \quad (2.6)$$

$$E_{ox}^E = \frac{V_D - V_{FG}}{t_{ox}} = \frac{V_D(1 - k_D) - \frac{Q_{FG}^E(t)}{C_T} - V_{SG} - V_{depl}}{t_{ox}} \quad (2.7)$$

Die Programmierung und Löschung der Speicherzelle erfolgt über ein zeitabhängiges Spannungssignal, das am Control-Gate bzw. Drainanschluss anliegt und folgende Form besitzt:

$$V_{P/E}(t) = \begin{cases} t \cdot \frac{V_{max}}{t_r} & : t \leq t_r \\ V_{P/E} = V_{max} & : t_r \leq t \leq t_h + t_f \\ V_{P/E} - t \cdot \frac{V_{max}}{t_f} & : t_p \geq t \geq t_p - t_f \end{cases} \quad (2.8)$$

Der Signalverlauf kann in Anstiegsflanke, Plateaubereich und Abstiegsflanke unterteilt werden. Im Bereich der Anstiegs- und Abstiegsflanke ändert sich die Spannung linear, wohingegen die Spannung im Plateaubereich dem Maximalwert V_{max} entspricht. Die

2 Grundlagen zu Speicherzellen

Pulslänge setzt sich anteilig aus diesen drei Bereichen zusammen: $t_p = t_r + t_h + t_f$, wobei t_r als Anstiegszeit, t_h als Haltezeit und t_f als Abstiegszeit bezeichnet wird. Die Spannungsänderung pro Zeiteinheit $a_r = V_{max}/t_r$ (Anstiegsflanke) und $a_f = -V_{max}/t_f$ (Abstiegsflanke) wird als Flankensteilheit bezeichnet.

Aus diesen Angaben kann die Floating-Gate-Ladung berechnet werden, indem der Fowler-Nordheim-Strom unter Verwendung der elektrischen Felder Gl.2.6 und Gl.2.7 integriert wird.

$$Q_{FG}^{P/E}(t) = \int_0^{t_p} I_{FN}(t) dt = A \int_0^{t_p} (E^{P/E})^2(t) e^{-\frac{B}{E^{P/E}(t)} dt} \quad (2.9)$$

Hiermit lässt sich die Änderung der Schwellspannung im programmierten und gelöschten Zustand direkt berechnen. Aus der Gl.2.4 folgt:

$$V_{CG} = 0 \Rightarrow V_{FG} = -Q_{FG}/C_T$$

Andererseits kann das Floating-Gate-Potenzial über den Control-Gate-Kopplungsfaktor berechnet werden:

$$V_{FG} = v_{th} k_{CG}$$

Aus diesen Gleichungen folgt ein Zusammenhang zwischen der Schwellspannungsänderung und der Änderung des Floating-Gate-Potenzials:

$$\Delta V_{FG} = \Delta Q_{FG}/C_T \text{ und } \Delta V_{FG} = \Delta v_{th} k_{CG} \Rightarrow \Delta v_{th} = \Delta Q_{FG}/C_{CG}$$

Wird die Schwellspannungsänderung im programmierten (gelöschten) Zustand bezogen auf die Schwellspannung im ladungsfreien Zustand (v_{thN}) definiert :

$$\Delta v_{thH} = v_{thH} - v_{thN} \text{ (programmiert)} \quad \Delta v_{thL} = v_{thL} - v_{thN} \text{ (gelöscht),}$$

so ergeben sich folgende Gleichungen:

$$a) \quad \boxed{v_{thL} = v_{thN} - \frac{Q_{FG}^E}{C_{CG}}} \quad b) \quad \boxed{v_{thH} = v_{thN} + \frac{Q_{FG}^P}{C_{CG}}} \quad (2.10)$$

2.4 Leitungsmechanismen in Isolatoren

In Isolatoren findet elektrische Leitung durch quantenmechanisches Tunneln oder durch thermische Anregung von Ladungsträgern statt. Im Fall hoher Temperaturen kann die Potenzialbarriere durch Schottky-Emission [29] überwunden werden. Ein weiterer Leitungsmechanismus ist die Pool-Frenkel-Emission [29], bei der Ladungsträger durch thermische Anregung entlang mehrerer Defektniveaus ins Leitungsband des Isolators gelangen. Bei den rein quantenmechanischen Leitungsmechanismen muss zwischen direktem Tunneln und Fowler-Nordheim-Tunneln unterschieden werden. Beim direkten Tunneln, das bei Schichtdicken kleiner 5 nm vorherrschend ist, wird die gesamte Isolatorschicht durchtunnelt, Bild 2.11 a), wohingegen bei Schichtdicken größer 5 nm das FN-Tunneln überwiegt, Bild 2.11 b). Hierbei wird die Isolatorschicht nur partiell durchtunnelt, da die Tunnelstrecke x_T kleiner als die Oxiddicke t_{ox} ist. Aufgrund der Spannungsabhängigkeit der Tunnelstrecke x_T weist das Fowler-Nordheim-Tunneln eine stärkere Feldabhängigkeit auf als das direkte Tunneln. Beide Tunnelmechanismen sind durch eine schwache Temperaturabhängigkeit charakterisiert. Weitere Leitungsmechanismen, die auf quantenmechanischem Tunneln basieren, sind das Trap assisted

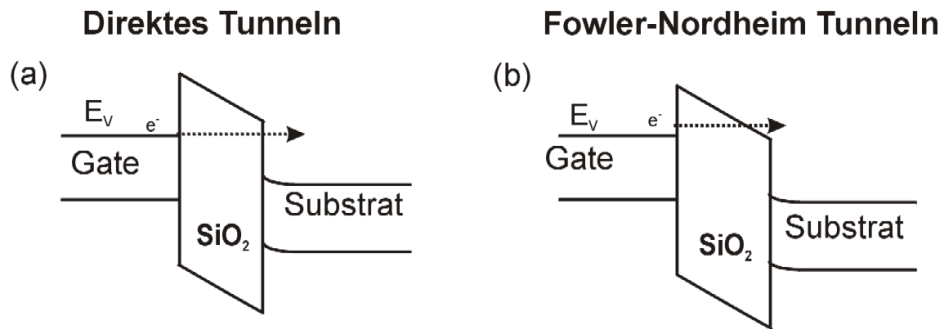


Bild 2.11: a) Das Bänderdiagramm zeigt den Bandverlauf beim direkten Tunneln. Die SiO_2 -Schicht wird hierbei vollständig durchtunnelt. b) Bei der Fowler-Nordheim Injektion wird die SiO_2 -Schicht nur partiell durchtunnelt.

Tunneling (TAT) [30,31] und das Phonon-assisted Tunneling [32,33,34]. Beim TAT erfolgt der Tunnelprozess über mehrere Defektniveaus innerhalb der SiO_2 -Barriere. Das Phonon-assisted Tunneling kombiniert das TAT mit der Anregung der Ladungsträger durch Phononen.

2.4.1 Theorie der Fowler-Nordheim-Leitung

Der Fowler-Nordheim-Strom nimmt in dieser Arbeit eine besondere Rolle ein und soll daher eingehend diskutiert werden.

Der von R. H. Fowler und L. Nordheim [35] untersuchte Leitungsmechanismus in Siliziumdioxid lässt sich als eine feldunterstützte quantenmechanische Emission von Ladungsträgern interpretieren. Die feldunterstützte Emission von Ladungsträgern kann mithilfe des Bänderdiagrammes in Bild 2.12 veranschaulicht werden. Der Potenzial-

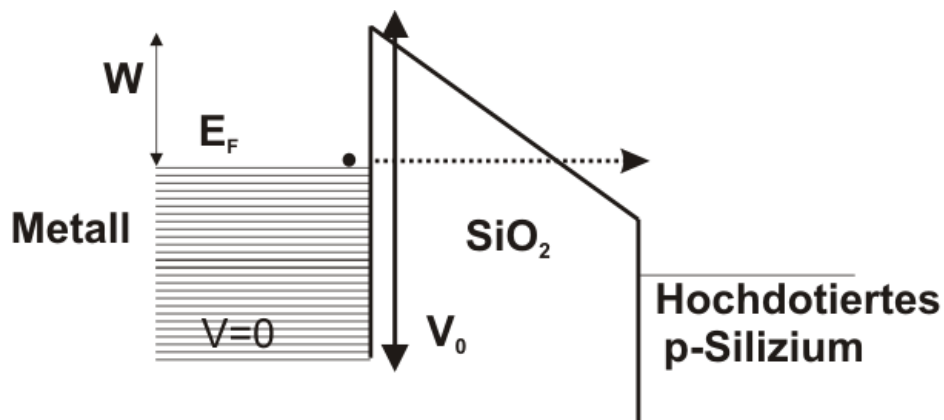


Bild 2.12: Darstellung des Bänderdiagrammes einer MOS-Struktur bei negativer Gate-Spannung.

wall zwischen Metall (links) und Silizium (rechts) wird durch Anlegen eines elektrischen Feldes abgesenkt. Hierdurch erhöht sich die Tunnelwahrscheinlichkeit und somit die Durchlässigkeit des Potenzialwalls für die Ladungsträger. Der hierdurch her-

2 Grundlagen zu Speicherzellen

vorgeschalteter Stromfluss wird durch das Fowler-Nordheim-Gesetz Gl.2.1 beschrieben.

$$j_{FN} = AE^2 e^{-\frac{B}{E}} \quad (\text{Fowler-Nordheim-Gesetz})$$

Hierbei steht j_{FN} für die Fowler-Nordheim-Stromdichte, E für das elektrische Feld und die Größen A , B werden als Fowler-Nordheim-Tunnelkonstanten bezeichnet. Für die Herleitung der Fowler-Nordheim-Gleichung wird angenommen, dass im Inneren des Metalls das Potential $V=0$ ist und an seiner Oberfläche ein Potentialsprung auf V_0 erfolgt. Bei angelegtem elektrischem Feld besitzt der Potenzialverlauf eine dreieckige Form:

$$V(z) = V_0 - eEz \quad (2.11)$$

Die folgende Berechnung der Stromdichte basiert auf der Bestimmung der Ladungsträger pro Energieintervall dE und deren Tunnelwahrscheinlichkeit bzw. Durchlässigkeit $D(E)$. Die Integration über alle Energiezustände liefert den Tunnelstrom. Ausgangspunkt für die Berechnung des Tunnelstromes ist folgende allgemeine Formel für die Stromdichte j .

$$j = e \int D v_z dn = e \int D v_z \frac{dp_x dp_y dp_z}{h^3} \quad (2.12)$$

Die Größe $dn = dp_x dp_y dp_z / h^3$ zählt die Zustände im Phasenraumelement $dp_x dp_y dp_z$, v_z ist die Elektronengeschwindigkeit in Richtung des Ziehfeldes und D ist die Durchlässigkeit des Potenzialwalls. Die Durchlässigkeit hängt von der Elektronenenergie ($\epsilon_z = p_z^2 / 2m_e$) in z -Richtung ab; m_e steht dabei für die Elektronenmasse im Metall und p_z bezeichnet die z -Komponente des Elektronenimpulses.

Die Integration in x , y -Richtung kann sofort ausgeführt werden. In Polarkoordinaten wird $dp_x dp_y$ zu $\rho d\rho d\phi$. Für das Integrationsintervall gilt:

$$0 < \rho < \sqrt{2m_e \epsilon_F - p_z^2} \quad (2.13)$$

Hiermit folgt:

$$\int dp_x dp_y = 2m_e \pi (\epsilon_F - E_z) \quad (2.14)$$

Wird dieser Ausdruck in 2.12 eingesetzt, so folgt:

$$j = \frac{4\pi e}{h^3} \int_0^{\sqrt{2m_e \epsilon_F}} (\epsilon_F - E_z) D p_z dp_z \quad (2.15)$$

Mit der Hilfsgröße $\epsilon = \epsilon_F - E_z$ ergibt sich:

$$j = \frac{4\pi m_e}{h^3} \int_0^{\epsilon_F} \epsilon D(\epsilon) d\epsilon \quad (2.16)$$

Die Durchlässigkeit des Potenzialwalls errechnet sich mithilfe der WKB-Näherung [36] zu:

$$D = \exp\left(-2\sqrt{\frac{2m_{ox}}{\hbar^2}} \int_{z_1}^{z_2} \sqrt{V(z) - E_z} dz\right) \quad (2.17)$$

Die Integrationsgrenzen z_1 und z_2 sind als klassische Umkehrpunkte bekannt und folgen aus der Bedingung $V(z) = E_z$. Anschaulich entspricht ihre Differenz der Tunnelstrecke, die ein Maß für die Tunnelwahrscheinlichkeit ist. Innerhalb der SiO_2 -Schicht wird an-

statt der Elektronenmasse m_e die effektive Masse m_{ox} verwendet. Im konkreten Fall eines linearen elektrischen Feldes führt die Berechnung der Durchlässigkeit auf eine elementare Integration (Gl. 2.18), die direkt ausführbar ist.

$$\int_{z_1}^{z_2} \sqrt{V(z) - E_z} dz = \int_{z_1}^{z_2} \sqrt{V_0 - eEz - \varepsilon_F + \varepsilon} dz \quad (2.18)$$

und Gl.2.19 liefert:

$$D(E_z) = e^{(-\frac{4\sqrt{2m_{ox}}}{3heE}(V_0 - \varepsilon_F + \varepsilon)^{3/2})} \quad (2.19)$$

Kombiniert man die Gleichungen 2.16 und 2.19 miteinander und führt ferner die Austrittsarbeit $\Phi = V_0 - \varepsilon_F$ ein, so ergibt sich:

$$j = \frac{4\pi em_e \Phi^2}{h^3} \int_0^{\varepsilon_F} e^{-\frac{4\sqrt{2m_{ox}}}{3heE} \Phi^{3/2} \varepsilon (1 + (\varepsilon/\Phi))^{3/2}} d\varepsilon \quad (2.20)$$

Die Integration des Ausdruckes Gl.2.20 ist in geschlossener Form nicht möglich. Es müssen daher einige Näherungen zu seiner Vereinfachung gemacht werden [36], die schließlich das Fowler-Nordheim-Gesetz liefert.

$$j_{FN} = \left[\frac{e^3 m_e}{8\pi m_{ox} \Phi} \right] E^2 e^{-\left[\frac{4\sqrt{2m_{ox}} \Phi^3}{3he} \right] \frac{1}{E}} \quad (2.21)$$

Der Vergleich von Gl. 2.21 mit Gl.2.1 liefert für die Tunnelkonstanten folgende Ausdrücke:

$$A = \frac{e^3 m_e}{8\pi m_{ox} \Phi} \quad B = \frac{4\sqrt{2m_{ox}} \Phi^3}{3he} \quad (2.22)$$

2.4.2 Temperaturabhängigkeit

Bei höheren Temperaturen kommt es zu Abweichungen von der klassischen Fowler-Nordheim-Gleichung, da die Barrierenhöhe Φ mit zunehmender Temperatur abnimmt und die FN-Injektion steigt [37,38]. Dieser Effekt ist primär auf die Änderung des Ferminiveaus zurückzuführen, wodurch die Barrierenhöhe Φ sinkt.

$$\Phi(T) = \Phi_0 + E_F(T) \quad (2.23)$$

$$j_{FN}(E, T) = j_{FN}(E) \frac{c\pi k_B T}{\sin(c\pi k_B T)} \quad \text{mit} \quad c = \frac{2\sqrt{2m_{ox}} \Phi}{ehE} \quad (2.24)$$

Hierbei steht $j_{FN}(E)$ für den temperaturunabhängigen FN-Strom, k_B ist die Boltzmann-Konstante, T bezeichnet die Temperatur und Φ steht für die Barrierenhöhe. In experimentellen Untersuchungen wird neben dem Fermi-niveau auch die Temperaturabhängigkeit von Φ_0 durch einen linearen Zusammenhang modelliert [39,79].

$$\Phi(T) = a + bT + \varepsilon_F(T), \quad (2.25)$$

mit a und b als Fitparameter.

2 Grundlagen zu Speicherzellen

2.4.3 Fowler-Nordheim-Leitung mit Oxidladungen

Die elektrische Beanspruchung von Speicherzellen verursacht im Tunneloxid Defekte und insbesondere Oxidladungen, die durch ihre elektrischen Felder die Fowler-Nordheim-Injektion nachhaltig verändern. Der Einfluss von Oxidladungen auf den Bandverlauf einer $\text{Si}/\text{SiO}_2/\text{Si}$ -Grenzfläche ist in Bild 2.13 dargestellt und zeigt positive und negative Oxidladungen mit ihren Schwerpunkten x_- und x_+ . Die Tunnelstrecke kann in die beiden Teilabschnitte 1 und 2 unterteilt werden, für die separat die Tunnelwahrscheinlichkeiten P_1 und P_2 sowie die elektrischen Felder E_1 und E_2 berechnet werden können. Es zeigt sich, dass durch negative Oxidladungen der Bandverlauf flacher wird und sich

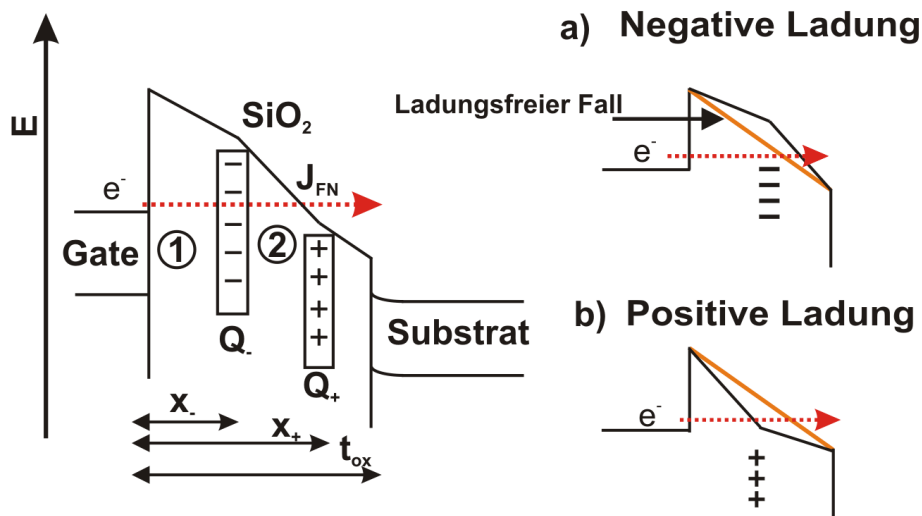


Bild 2.13: Der Bandverlauf einer MOS-Struktur bei negativer Gatespannung. Deutlich ist die Bandverbiegung durch die Oxidladungen Q_+ und Q_- zu erkennen, deren Schwerpunkte sich im Abstand x_+ und x_- von der SiO_2 -Oberfläche befinden. a) Verkürzung der Tunnelstrecke durch negative Oxidladungen. b) In Gegenwart von positiven Oxidladungen findet eine Verlängerung der Tunnelstrecke statt.

die Tunnelstrecke gegenüber dem ladungsfreien Oxid verlängert, Bild 2.13 a). Positive Oxidladungen verursachen hingegen einen steileren Bandverlauf und verkürzen die Tunnelstrecke Bild 2.13 b), wodurch der FN-Strom zunimmt. Wie im ladungsfreien Fall ergeben sich die Tunnelwahrscheinlichkeiten für die Abschnitte 1 und 2 aus der WKB-Näherung, wobei dreieckige Potenzialverläufe zugrunde gelegt werden. Die elektrischen Felder werden über die Poission-Gleichung und den gaußschen Satz bestimmt. Da die Ladungsträger die Bereiche 1 und 2 durchtunneln müssen, wird für die Tunnelwahrscheinlichkeit und damit für den Tunnelstrom folgender Produktansatz gemacht [40]:

$$j_{FN} = CE_1^2 P_1 P_2, \quad (2.26)$$

wobei E_1 das elektrische Feld in Abschnitt 1 bezeichnet und C ein Proportionalitätsfaktor ist. Diese Gesetzmäßigkeit geht in das bekannt Fowler-Nordheim-Gesetz über, wenn der Tunnelvorgang durch den zweiten Teilabschnitt vernachlässigbar ist.

2.4.4 Modell mit Ladungseinbau

Das im vorangegangenen Kapitel vorgestellte Modell vereinfacht sich deutlich, wenn nur der Einfang negativer Ladungsträger betrachtet wird. Es gibt verschiedene Theorien, die die Fowler-Nordheim-Injektion in Abhängigkeit von Oxidladungen modellieren [41,42,43]. Das von Papadas et al. [42] entwickelte Modell soll im Folgenden näher betrachtet werden und später bei der Modellierung von Speicherzellen eingesetzt werden. Dieses Modell hält an der Feldabhängigkeit des klassischen Fowler-Nordheim-Gesetzes fest und berechnet die elektrischen Felder unter Beachtung der Oxidladungen, Bild 2.14 a). Diese Annahme setzt voraus, dass der Tunnelstrom primär durch den Be-

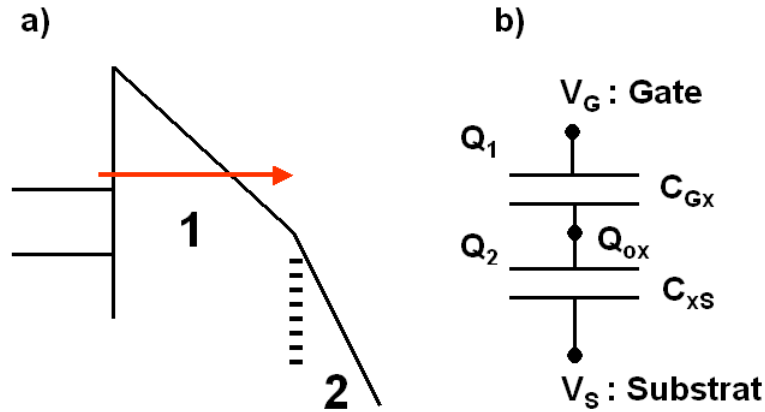


Bild 2.14: a) Bänderdiagramm mit Oxidladung. Wenn der Tunnelstrom durch den zweiten Teilabschnitt vernachlässigbar ist, kann das Fowler-Nordheim-Gesetz angewendet werden. b) Ersatzschaltbild für das Tunneloxid mit Oxidladung. Die Oxidladung verursacht einen Spannungsabfall, der über einen kapazitiven Spannungsteiler modelliert werden kann.

reich 1 erfolgt und der Stromfluss durch den zweiten Teilabschnitt vernachlässigbar ist. Der durch die Oxidladungen verursachte Spannungsabfall lässt sich über einen kapazitiven Spannungsteiler bestimmen, Bild 2.14 b). Das Ersatzschaltbild für das gealterte Oxid besteht aus zwei in Reihe geschalteten Kapazitäten C_{Gx} und C_{xs} mit den Ladungen Q_1 und Q_2 . Aus dem Ladungserhaltungssatz folgt unter Anwendung des Kondensatorgesetzes die Gleichung 2.27. Hieraus ergeben sich die elektrischen Felder für die Fowler-Nordheim-Injektion, Gl.2.28

$$V_{ox} - V_S = \frac{Q_{ox}}{C_t} + \frac{C_{Gx}}{C_t}(V_G - V_S) \quad \text{mit} \quad C_t = C_{Gx} + C_{xs}, \quad (2.27)$$

wobei V_{ox} für den Spannungsabfall über der jeweiligen Kapazität C_{Gx} und C_{xs} steht.

$$E = \begin{cases} E_S = \frac{V_S - V_{ox}}{t-x} & : V_S(t) > 0, V_G = GND \\ E_G = \frac{V_G - V_{ox}}{t-x} & : V_G(t) > 0, V_S = GND \end{cases} \quad (2.28)$$

2 Grundlagen zu Speicherzellen

Mit den Gleichungen 2.28 kann der Fowler-Nordheim-Strom über Gl.2.29 gefunden werden.

$$J_{FN} = A(E_{G/S})^2 e^{-\frac{B}{E_{G/S}}} \quad (2.29)$$

Mit diesen Gleichungen lässt sich der FN-Strom zeitabhängig und unter Berücksichtigung von Oxidladungen simulieren. Der Ladungsschwerpunkt wurde hierzu in der Oxidmitte angenommen und es wurden die Tunnelkonstanten $A=1 \cdot 10^{-6} \text{AV}^{-2}$ und $B=3 \cdot 10^{10} \text{mV}^{-1}$ gesetzt. Für die Oxiddicke wurde ein Wert von 9.4 nm gewählt.

Der verwendete Spannungspuls ist durch seine Anstiegszeit t_r , Abstiegszeit t_f und die Spannungsamplitude V_{max} charakterisiert, siehe Gl.2.30. Die sich hieraus ergebende Pulslänge t_p ergibt sich zu $t_p = t_r + t_f$.

$$V(t) = \begin{cases} t \cdot \frac{V_{max}}{t_r} & : t \leq t_r \\ V_{max} - t \cdot \frac{V_{max}}{t_f} & : t_p \geq t \geq t_r \end{cases} \quad (2.30)$$

In Bild 2.15 ist für den ladungsfreien Fall die FN-Stromdichte und die Gatespannung als Funktion der Zeit dargestellt. Das Maximum der Gatespannung liegt bei 14 V und fällt mit der maximalen Stromdichte zusammen. Die schraffierte Fläche entspricht der akkumulierten injizierten Ladung, die nach einem Spannungspuls geflossen ist; ihre Bestimmung erfolgte über die Integration des FN-Stromes. Der Einfluss der Oxidladung

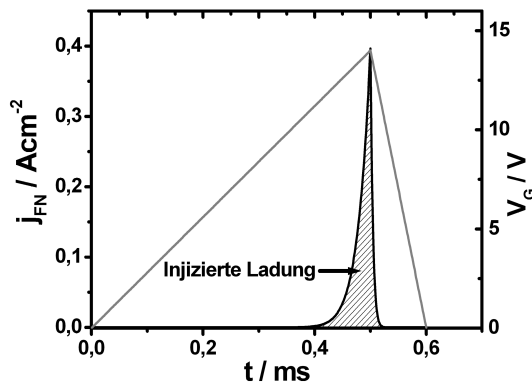


Bild 2.15: Die FN-Stromdichte ist als Funktion der Pulszeit dargestellt (linke Achse). Der Spannungspuls ist auf der rechten Achse zu sehen.

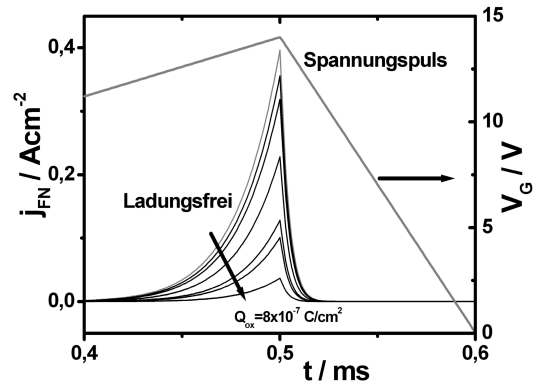


Bild 2.16: Die FN-Stromdichte wurde für verschiedene Oxidladungen zeitabhängig simuliert.

auf das Strommaximum bei 14 V ist in Bild 2.16 dargestellt und zeigt, dass sowohl die Höhe als auch die Lage des Strommaximums deutlich von der Oxidladung abhängen. Das Strommaximum in Abhängigkeit von der Oxidladung ist in Bild 2.17 zu sehen. Der FN-Strom wurde hierzu bei $V_G=14 \text{ V}$ berechnet und als Funktion der eingebauten Oxidladung dargestellt. Auch hier zeigt sich eine deutliche Abnahme des FN-Stromes.

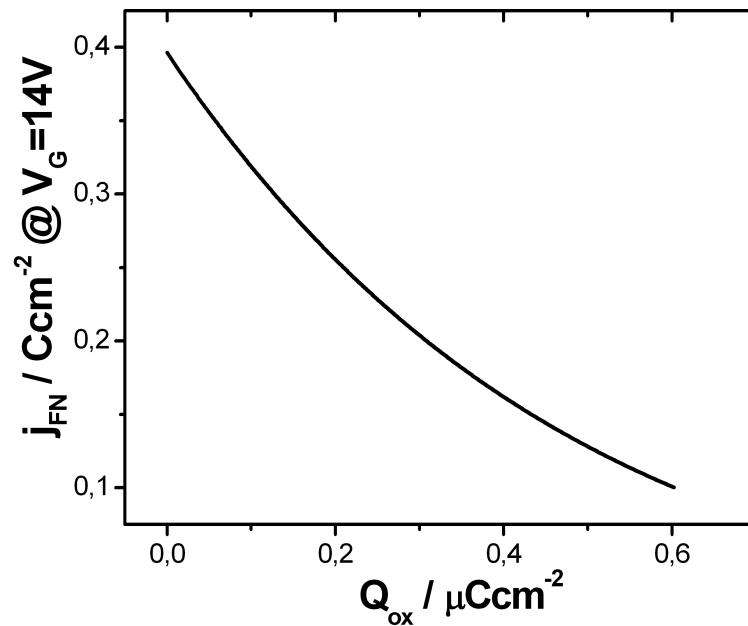


Bild 2.17: Die FN-Stromdichte wurde bei einer Gatespannung von 14 V berechnet und in Abhängigkeit von der Oxidladungen dargestellt.

2.5 Defekte in Siliziumdioxid

Nach den Erläuterungen zum FN-Strom soll nachfolgend näher auf Oxiddefekte, Oxidladungen und ihre Entstehungsursachen eingegangen werden.

2.5.1 Defekt-Klassifizierung

Defekte im Siliziumdioxid können bezüglich ihrer räumlichen Lage im Oxid und ihrer Entstehungsursache differenziert werden. So unterteilt man die Defekte in Oberflächendefekte (Interface-Traps), grenzflächennahe Defekte (Near-Interface-Traps oder Border-Traps) und in Oxiddefekte, die im Inneren des Oxides angesiedelt sind, Bild 2.18. Die Border-Traps [44] unterscheiden sich strukturell unwesentlich von den Oberflächendefekten, sind maximal 3 nm von der Grenzfläche entfernt und können über direktes Tunneln mit dem Substrat- oder Gatekontakt in Wechselwirkung treten [45,49]. Die Grenzfläche zwischen Silizium und thermisch oxidiertem Silizium (gateseitig und substratseitig) erstreckt sich über etwa 2 Atomlagen [49]. Die Defekte in diesen Bereichen können in zwei Klassen unterteilt werden [46,47]: Die einen werden als Dangling-Bond-Defekte bezeichnet und auf fehlende Bindungspartner des Siliziumatoms zurückgeführt. Die andere Defektart wird durch gestreckte *Si-Si*-Bindungen verursacht und als Tailstate U_T bezeichnet. Die Dangling-Bond-Defekte (U_M -, P_{b0} -, P_H und E_0 -Defekt) können in Bezug auf die Anzahl der rückgebundenen Sauerstoffatome unterschieden werden, Bild 2.19 a). Sämtliche Dangling-Bond-Defekte sind durch ein ungebundenes Si-Atom charakterisiert, das Elektronen aufnehmen und abgeben kann. Darüber hinaus unterscheiden sich die Traps in ihrer energetischen Lage in Bezug auf das Valenz- bzw.

2 Grundlagen zu Speicherzellen

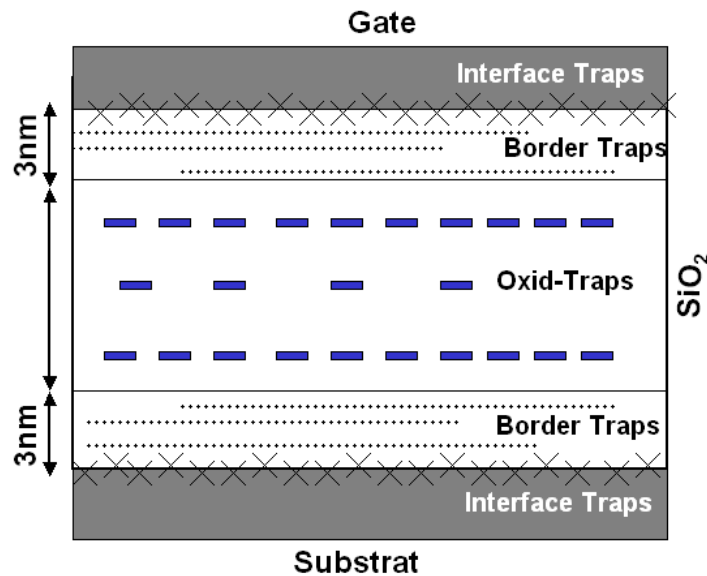


Bild 2.18: Aufteilung der Oxiddefekte in Grenzflächendefekte (Interface-Traps), grenzflächennahe Defekte (Border-Traps) und Oxiddefekte im Inneren des Oxides [49]

Leitungsband des Siliziums. In Bild 2.19 b) sind die Zustandsdichten der verschiedenen Defektarten in Abhängigkeit von der Defektenergie dargestellt [46,47], wobei die Leitungsband- und Valenzbandkante des Siliziums als Referenzpunkte angegeben sind. Es zeigt sich, dass die P_L -Defekte in der unteren Hälfte der Bandlücke und die P_M -Defekte im oberen Bereich der Bandlücke zu finden sind, wohingegen die U_M -Defekte innerhalb der gesamten Bandlücke auftreten können. Die U_M -Defekte besitzen oberhalb des intrinsischen Fermi-niveaus E_i ein donatorartiges Verhalten und für Energien unterhalb von E_i ein akzeptorartiges Verhalten. Die P_M - und P_L -Defekte zeigen hingegen in beiden Fällen ein donatorartiges Verhalten [46]. Die E' -Defekte liegen bereits außerhalb der Si-Bandlücke und können dem Oxidbereich zugeordnet werden. Die Ursache für den E_0 -Defekt ist ein fehlendes Sauerstoffatom bzw. eine Sauerstoffleerstelle in der SiO_2 -Matrix.

In Gegenwart von Wasserstoff kann sich an die Sauerstoffleerstelle ein Wasserstoffatom anlagern und eine Wasserstoffbrücke ausbilden. Durch elektrischen Stress können hieraus verschiedene Defektkonfigurationen hervorgehen. Mögliche Defektkonfigurationen sind die E'_β - und E'_4 -Defekte, die aus strukturellen Änderungen der Wasserstoffbrücke hervorgehen. Diese Unterschiede resultieren aus der räumlichen Orientierung des freien Valenzelektronenorbitals und aus unterschiedlichen Polaritäten der Bindung zwischen dem Siliziumatom und dem SiH-Komplex [48,49]. Der für die Entstehung dieser Defektart erforderliche Wasserstoff befindet sich im Siliziumdioxid in ausreichender Menge, da z. B. das Tempern in Wasserstoffatmosphäre (Post Annealing) durchgeführt wird [50]. Dieser wasserstoffinduzierte Defekttyp spielt insbesondere bei der Erklärung des stressinduzierten Leckstromes SILC (Stress Induced Leakage Current) eine wichtige Rolle [51]. Die Ursachen der stressinduzierten Leckströme sollen im nächsten Kapitel behandelt werden.

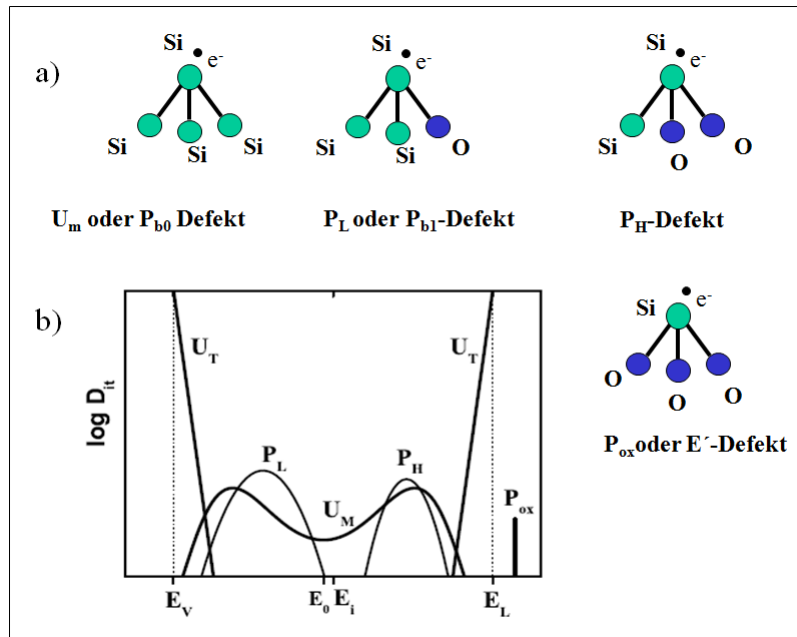


Bild 2.19: a) Schematische Darstellung der atomaren Struktur von Oberflächen- und Oxiddefekten (U_M , P_{b1} , P_H , E') [46]. b) Zustandsdichte verschiedener Defekttypen in Abhängigkeit von der Energie [46,47].

2.5.2 Isolatorschädigungsmodell

Die Entstehung von Defekten in Siliziumdioxid ist ein komplexer Vorgang, für dessen Beschreibung verschiedene Modelle entwickelt worden sind. Die drei wichtigsten sollen hier kurz vorgestellt werden, siehe Bild 2.20.

(1) Hydrogen-Release: Dieser Schädigungsmechanismus basiert auf der anodenseitigen Freisetzung von Wasserstoffspezies durch elektrischen Stress. Die freigesetzten Spezies diffundieren durch das Oxid und brechen durch Wechselwirkungen SiO-Bindungen auf [52,53].

(2) Anode-Hole-Injection: Bei diesem Model wird von einer anodenseitigen Entstehung von Elektron-Lochpaaren ausgegangen. Die hierfür erforderliche Energie liefern Elektronen, die anodenseitig ihre Energie über Relaxationsprozesse abgeben. Die dadurch entstandenen Löcher tunneln durch die Potenzialbarriere und können so im SiO_2 Bindungen aufbrechen [54,55].

(3) Impact-Ionization: Bei dickeren Oxiden > 20 nm spielt die Erzeugung von Löchern ebenfalls eine zentrale Rolle bei der Oxidschädigung. Die Elektronen können bei ihrer Bewegung in Richtung Kathode Energie aufnehmen und diese durch Streuprozesse mit Phononen und über die Erzeugung von Elektron-Lochpaare abgeben [52,56]. Dabei bewegen sich die Löcher in Richtung Kathode und schädigen das Oxid. Für dünne Oxide ist dieser Schädigungsmechanismus von geringer Bedeutung, da die Energien für die Generierung von Elektron-Loch-Paaren nicht ausreichen. Die Erzeugung von Defektniveaus spielt insbesondere in der Percolationstheorie bei der Erklärung von

2 Grundlagen zu Speicherzellen

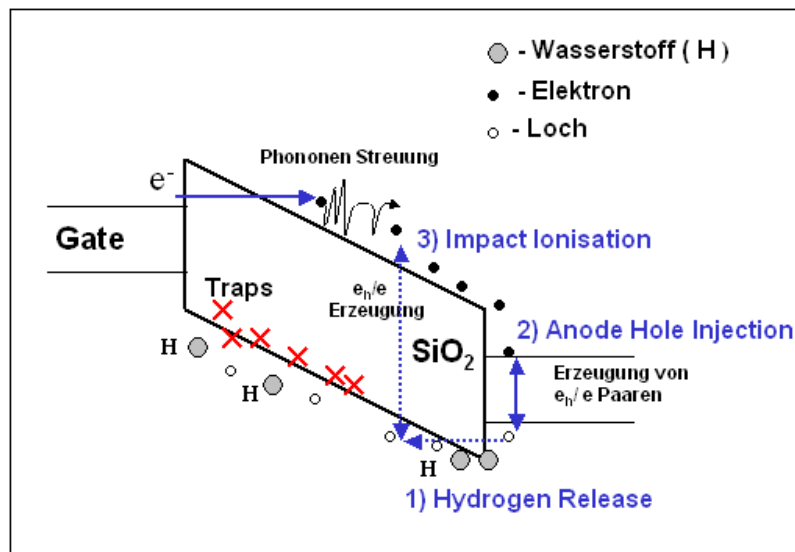


Bild 2.20: Schematische Darstellung der wichtigsten Schädigungsmechanismen in SiO_2 (1) Freisetzung von Wasserstoffspezies (Hydrogen-Release) (2) Anodenseitige Löcherinjektion (Anode-Hole-Injection) (3) Stoßionisation und Erzeugung von Elektron-Lochpaaren (Impact-Ionisation)[49]

Oxiddurchbrüchen und den dazugehörigen Statistiken eine zentrale Rolle [57,58]. Die Grundannahme hierfür ist die zufällige Erzeugung von Defekten innerhalb des gesamten Oxidvolumens. Die Defektniveaus besitzen eine räumliche Ausdehnung von ca. 0.9 nm [49] und bilden bei einem Oxiddurchbruch einen zusammenhängenden, leitfähigen Pfad aus, der die SiO_2 -Schicht überbrückt, Bild 2.21 a) Entlang des Defektpfades können sich die Elektronen durch direktes Tunneln bewegen.

Ein Oxiddurchbruch ist exemplarisch in Bild 2.21 b) zu sehen. An einer großflächigen Kapazität wurde gateseitig eine Spannung von 13.5 V angelegt und der korrespondierende Gatestrom zeitabhängig gemessen. Ein kontinuierlich abnehmender Gatestrom kann festgestellt werden. Aufgrund des Stromflusses werden Oxidladungen erzeugt, die das elektrische Feld über dem Tunneloxid verringern. Nach etwa 80 s bricht das Oxid durch, was am sprunghaften Anstieg des Gatestromes erkennbar ist. In der Praxis wird der Oxiddurchbruch über die flächennormierte Durchbruchladung Q_{BD} (Charge to Breakdown) charakterisiert, die aus dem geflossenen Strom I_s , der Zeit bis zum Durchbruch t_{BD} und der spezifischen Fläche A_{inj} berechnet wird.

$$Q_{BD} = \frac{1}{A_{inj}} \int_0^{t_{BD}} I_s(t) dt \quad (2.31)$$

Eine hohe Durchbruchladung bzw. Durchbruchzeit korrespondiert mit einer hohen Lebensdauer des betrachteten Oxides und vice versa impliziert eine geringe Durchbruchzeit und -ladung eine kurze Lebensdauer des Oxides. Ein weiteres Phänomen, das durch Defektstellen verursacht wird, ist der Stress Induced Leakage Current (SILC). Durch elektrischen Stress entstehen innerhalb der Oxidschicht Defekte, die Zwischen-

position in einem mehrstufigen Tunnelprozess sind. Die Abstände zwischen den De-

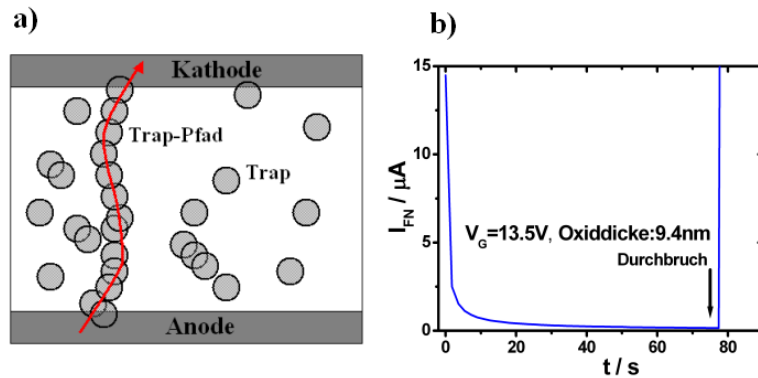


Bild 2.21: a) Skizze zum Percolations-Modell: Ein Oxiddurchbruch liegt vor, wenn ein aus Defektstellen bestehender Pfad die Oxidschicht überbrückt. b) Oxiddurchbruch: An einer Kapazität wurde gateseitig eine Spannung von 13.5 V angelegt und der Gatestrom gemessen. Nach etwa 75 s bricht das Oxid durch.

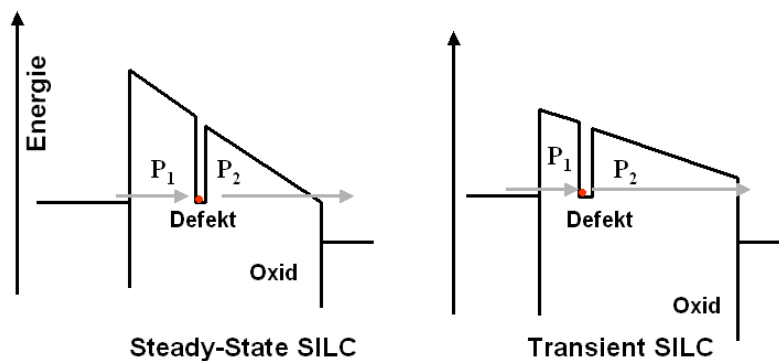


Bild 2.22: a) Steady-State SILC: Die Besetzung und Entleerung des Defektzentrums ist etwa gleichwahrscheinlich, so dass ein zeitunabhängiger Leckstrom fließen kann. b) Transienter-SILC: Die Besetzung der Defektstelle ist wahrscheinlicher als seine Entleerung ($P_1 \gg P_2$).[49]

fektniveaus liegen zwischen 3 und 4 nm und können von Ladungsträgern durch direktes Tunneln zurückgelegt werden. Die so verursachten Leckströme hängen vor allem von den stressinduzierten Oxiddefekten ab und werden daher als stressinduzierte Leckströme bezeichnet. In einem zweistufigen Tunnelprozess mit einem Defektzentrum als Zwischenposition muss zwischen dem „Transient“-SILC und dem „Steady State“-SILC unterschieden werden [49]. Die Leckströme hängen in diesem speziellen Fall vom Produkt der Tunnelwahrscheinlichkeiten $P_1 \cdot P_2$ ab und werden durch die jeweils kleinste Tunnelwahrscheinlichkeit begrenzt. Beim „Steady State“-SILC ist die Besetzungs- und Entleerungsrate der Defektniveaus ausgeglichen, wodurch ein zeitunabhängiger Leckstrom fließen kann, Bild 2.22 a). Wenn die Besetzung von Defektniveaus mit deutlich größerer Wahrscheinlichkeit stattfindet als ihre Entleerung ($P_2 \ll P_1$), so kann kurzzeitig ein

2 Grundlagen zu Speicherzellen

zeitabhängiger Leckstrom fließen, der als „Transient“-SILC bezeichnet wird, Bild 2.22 b).

2.6 Oxidladungen

Oxidladungen in Siliziumdioxid können nach E. B. Deal [59] hinsichtlich ihrer Position im Siliziumoxid und ihrer Entstehungsursache unterschieden werden, Bild 2.23. An der Grenzfläche zwischen Silizium und Siliziumdioxid befinden sich die grenzflächennahen Oxidladungen Q_{ss} , die durch die Besetzung von Oberflächendefekten entstehen (U_M , P_L , P_H - Defekte). Räumlich über den oberflächennahen Oxidladungen sind feste Oxidladungen Q_F lokalisiert. Die Oxidladungen Q_{ox} können im gesamten Volumen angetroffen werden und sind auf Defekte zurückzuführen, die von Elektronen oder Löchern besetzt werden. Mögliche Ursachen für Oxiddefekte sind Stress durch elektrische Beanspruchung oder durch die Halbleiterprozessierung, wie z. B. Reinigungs- und Oxidationsprozesse.

Des Weiteren gibt es freibewegliche Alkaliionen Q_M , wie z. B. Li^+ (Lithium), K^+ (Kalium), Na^+ (Natrium), die sich aufgrund ihrer Polarität im gatenahen Bereich ansammeln. Dieser Ladungstyp ist auf Kontaminationsprobleme zurückzuführen und tritt in den heutigen Technologien nur noch selten auf.

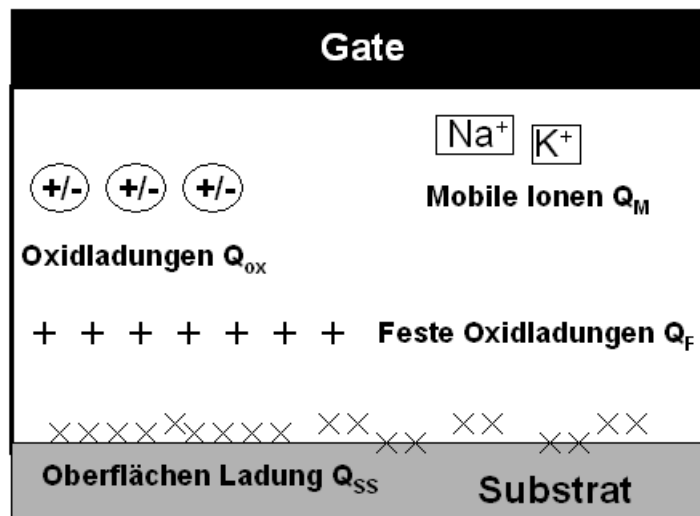


Bild 2.23: Klassifikation von Oxidladung nach B. E. Deal [59]. Es wird zwischen grenzflächennahen Oxidladungen, festen Oxidladungen und Volumen-Oxidladungen unterschieden. Die beweglichen Oxidladungen bestehen aus Alkaliionen.

2.6.1 Ladungseinbau: Kinetische Modelle

Für die Modellierung des Ladungseinbaues werden nachfolgend mathematische Modelle vorgestellt, die den Einbau von Oxidladungen in Abhängigkeit von der injizierten

Ladung Q_{inj} beschreiben. Die injizierte Ladung ist ein Maß für den Degradationsgrad des Oxides und kann aus der Stromdichte durch Integration über die Stressdauer berechnet werden.

$$Q_{inj} = \int_0^{T_s} j_s dt \quad (2.32)$$

Hierbei ist T_s die Stressdauer und j_s die Stromdichte. Das einfachste Modell zur Beschreibung des Ladungseinbaues basiert auf einer Ratengleichung erster Ordnung („First Order“-Modell) [60-62] und lässt sich mit folgenden Annahmen herleiten.

1. Das Siliziumoxid weist eine anfängliche Defektdichte N_{it} auf, die während der Halbleiterherstellung und Prozessierung entstanden ist. Eine genaue Kenntnis der Defektverteilung innerhalb der Oxidschicht ist nicht erforderlich. Die Generierung von neuen Defekten und eine mögliche Entleerung von bereits besetzten Defektniveaus (Detrapping) wird nicht berücksichtigt.
2. Der Wirkungsquerschnitt für den Ladungseinfang wird mit σ bezeichnet und ist in diesem Modell eine zeit- und ortsunabhängige Oxidgröße. Sie ist ein Maß für die Einfangwahrscheinlichkeit eines Ladungsträgers durch ein Defektzentrum.
3. Die injizierte Ladung kann über Gl. 2.32 bestimmt werden.

Im Folgenden soll über die zeitliche Änderung der Oxidladung die Ratengleichung bestimmt werden, deren Lösung einen funktionalen Zusammenhang zwischen Oxidladung und injizierter Ladung liefert.

Im Oxid befinden sich Defektniveaus, die durch Elektronen besetzt werden können. Unter der Annahme, dass bereits n_{ox} Defekte mit Elektronen besetzt sind, gibt es $N_{it} - n_{ox}$ Möglichkeiten für den Ladungseinfang. Mithilfe des Wirkungsquerschnittes σ ergibt sich für die Besetzungswahrscheinlichkeit: $p = (N_{it} - n_{ox})\sigma$.

Wird mit der Anzahl der Ladungsträger pro Zeiteinheit ($j_s dt/q$) multipliziert, so folgt hieraus die zeitliche Änderung der Oxidladung $dQ_{ox} = q dn_{ox}$.

$$dn_{ox} = -(N_{it} - n_{ox}) \frac{\sigma j_s dt}{q} \quad \text{oder} \quad \frac{dQ_{ox}}{dt} = -q(N_{it} - n_{ox})\sigma j_s(t)/q \quad (2.33)$$

Diese Differenzialgleichung kann leicht integriert werden und ergibt

$$Q_{ox}(Q_{inj}) = qn_{ox} = qN_{it}(1 - e^{-\frac{\sigma Q_{inj}}{q}}) \quad \text{mit} \quad Q_{inj} = \int_0^{T_s} j_s dt \quad (2.34)$$

Für die praktische Anwendung musste dieses Gesetz modifiziert werden, da die Annahmen 1 bis 3 die Realität nur eingeschränkt wiedergeben. So wurde von M. S. Liang et al. [62] festgestellt, dass zusätzlich zu den anfänglich vorhandenen Defekten neue Defekte entstehen. Um die neu generierten Defekte berücksichtigen zu können, wurde eine konstante Defekterzeugungsrate g eingeführt (Gl. 2.35), wobei die Größe N_{gt} die neu generierten Defekte bezeichnet.

$$g = \frac{dN_{gt}}{dt} = \frac{j_s}{g} \frac{dN_{gt}}{dQ_{inj}} \quad \text{mit} \quad Q_{inj} = \frac{j_s}{q} t \quad (2.35)$$

2 Grundlagen zu Speicherzellen

Ferner wird ein linearer Zusammenhang zwischen injizierter Ladung und der Stresszeit angenommen. Neben dem Wirkungsquerschnitt für die anfänglich vorhandenen Defekte σ_i wird ein weiterer Wirkungsquerschnitt für die neu generierten Traps σ_g benötigt. Mit diesen Annahmen und Begriffen können Differenzialgleichungen analog zu Gl.2.33 formuliert werden, die als Lösung Gleichung 2.36 liefern:

$$Q_{ox}(Q_{inj}) = qn_{ox} = qN_{it}(1 - e^{-\frac{\sigma_i Q_{inj}}{q}}) + \frac{gq^2}{j_s}(Q_{inj} - \frac{1}{\sigma_g}(1 - e^{-\frac{\sigma_g Q_{inj}}{q}})) \quad (2.36)$$

Die Gleichung 2.36 besitzt einen linearen Anteil, der die Erzeugung von Defekten beschreibt und zwei Exponentialfunktionen zur Beschreibung der Defektbesetzung. Die Exponentialfunktionen hängen von den Wirkungsquerschnitten σ_i und σ_g ab, die die Geschwindigkeit der Defektbesetzung bestimmen und ihrerseits von der Stromdichte [62,63] und den Eigenschaften des verwendeten Dielektrikums abhängen.

Für hohe injizierte Ladungen geht Gl.2.36 in einen linearen Ausdruck über,

$$Q_{ox}(Q_{inj}) \longrightarrow N_{it}q + \frac{q^2 g}{j_s} Q_{inj} + \frac{q^2 g}{\sigma_g j_s}, \quad \text{für } \sigma Q_{inj}/q \gg 1$$

der zeigt, dass im Bereich hoher Ladungsinjektionen die neu generierten Defekte den Ladungseinbau bestimmen [63,133]. Die Wirkungsquerschnitte können quantenmechanisch über die bornsche Näherung [64] berechnet werden, wenn das Wechselwirkungspotenzial zwischen Elektron und Defektzentrum bekannt ist. In [65] wird ein Zusammenhang zwischen Wirkungsquerschnitt, Polarisierbarkeit α und der Permittivitätszahl ϵ des Dielektrikums angegeben, wobei für die Wechselwirkung zwischen Elektron und (neutralem) Defektzentrum ein Potenzial mit r^{-4} -Abhängigkeit angenommen wurde.

$$V(r) = -\frac{\alpha q^2}{2(4\pi\epsilon)^2} \frac{1}{r^4} \quad \text{Bornsche Näherung: } \Rightarrow \sigma \propto (\frac{\alpha}{\epsilon^2})^2$$

2.6.2 Ladungseinbau: Potenzgesetzmäßigkeit

Die nun folgende Modellierung wurde von Papadas et al. entwickelt [66-68] und zeigt, dass der Zusammenhang zwischen Oxidladung und injizierter Ladung über eine Potenzgesetzmäßigkeit beschreibbar ist. Auf empirischem Weg konnte gezeigt werden, dass die Ableitung der Oxidladung nach Q_{inj} proportional zu Q_{inj}^n ist, worüber die folgende Potenzgesetzmäßigkeit gefunden wurde.

$$\frac{dQ_{ox}}{dQ_{inj}} = K Q_{inj}^n \implies \boxed{Q_{ox} = \frac{K}{n+1} Q_{inj}^{n+1}} \quad (2.37)$$

Der Exponent n ist vom Schädigungsmechanismus abhängig und liegt zwischen -0.7 und -0.6, wohingegen der K -Wert vom Trapping-Yield abhängt, der ein Maß für das Verhältnis aus geflossener und der vom Oxid eingefangenen Ladung ist. Für die theoretische Herleitung des Potenzgesetzes wird von folgenden Annahmen ausgegangen:

1. Es wird vorausgesetzt, dass der Einfang von negativen Oxidladungen dominiert; positive Oxidladungen werden nicht berücksichtigt.

2. Für die Defektdichte N_{it} und den Wirkungsquerschnitt σ werden ortsabhängige Funktionen eingeführt, Gl 2.38 a) und 2.39 a). Beide Größen werden in Bezug auf die SiO_2 -Oberfläche angegeben und enthalten die charakteristischen Längen L_n und L_σ . Trägt nur eine Grenzfläche zum Ladungseinfang bei, so gelten die Einzelprofile 2.38 b) und 2.39 b).

$$a) \quad \frac{N_t(x)}{N_{it}} = e^{\frac{x}{L_n}} + e^{\frac{x-1}{L_n}} \quad b) \quad \frac{N_t(x)}{N_{it}} = e^{\frac{x}{L_n}} \quad (2.38)$$

$$a) \quad \frac{\sigma(x)}{\sigma_0} = e^{-\frac{x}{L_\sigma}} + e^{\frac{x-1}{L_\sigma}} \quad b) \quad \frac{\sigma(x)}{\sigma_0} = e^{-\frac{x}{L_\sigma}} \quad (2.39)$$

3. Es wird an der Grundgleichung des „First Order“-Modells festgehalten und die Ortsabhängigkeit von $N_t(x)$ und $\sigma(x)$ durch Integration über die Oxiddicke (t_{ox}) berücksichtigt.

$$\rho(x) = -N_t(x) \left(1 - \exp\left(-\frac{\sigma(x)Q_{inj}}{q}\right)\right) \quad (2.40)$$

$$Q_{ox}(Q_{inj}) = q \int_0^{t_{ox}} \rho(x) dx \quad (2.41)$$

Das Potenzgesetz soll hier unter vereinfachten Bedingungen abgeleitet werden und es soll nur eine Grenzfläche zum Ladungseinfang beitragen, sodass Gl.2.38b) und Gl.2.39 b) für $N(x)$ und $s(x)$ verwendet werden. Die Oxidladung kann formal über Gl.2.41 berechnet werden, indem Gl.2.40 unter Berücksichtigung von 2.38 b) und 2.39 b) in Gl.2.41 eingesetzt wird:

$$Q_{ox}(Q_{inj}) = \int_0^{t_{ox}} q N_{it} e^{\frac{x}{L_n}} [1 - e^{-(Q_{inj}\sigma_0 e^{-\frac{x}{L_\sigma}})}] dx \quad (2.42)$$

Dieser Ausdruck wird unter Verwendung folgender Abkürzungen umgeformt:

$$K^- = -\frac{\Delta Q_{ox}}{q N_{it} t_{ox}}, \quad \xi_n = \frac{L_n}{t_{ox}}, \quad \xi_\sigma = \frac{L_\sigma}{t_{ox}}, \quad \xi = \frac{x}{t_{ox}} \quad \text{und} \quad N_0 = \sigma_0 Q_{inj} / q.$$

Die Hilfsgröße K^- ist dimensionslos und misst die Oxidladung in Einheiten von $q N_{it} / t_{ox}$.

$$K^- = \int_0^1 e^{\frac{\xi}{\xi_n}} (1 - \exp(-N_0 e^{\frac{\xi}{\xi_\sigma}})) d\xi \quad (2.43)$$

Mit der Substitution $\Psi = N_0 e^{\frac{\xi}{\xi_\sigma}}$ folgt hieraus nach Ausführung der Integration:

$$K^- = \xi_n (e^{\frac{1}{\xi_n}} - 1) + \xi_\sigma [N_0^a \Gamma(-a, N_0) - N_0^a \Gamma(-a, N_0 e^{1/\xi_\sigma})] \quad (2.44)$$

Γ steht für die unvollständige Gammafunktion (Anhang in [66]), die von N_0 und damit von Q_{inj} abhängt. Die Konstante a ergibt sich aus den charakteristischen Konstanten L_n und L_σ ($a = L_n / L_\sigma$). Die Gammafunktionen können durch Reihenentwicklung unter Berücksichtigung der Näherungsannahmen $N_0 \gg 1$ und $N_0 e^{-\frac{1}{\xi_n}} \ll 1$ stark vereinfacht werden. Das Ergebnis der Umformungen liefert eine Potenzgesetzmäßigkeit zwischen

2 Grundlagen zu Speicherzellen

Oxidladung und injizierter Ladung.

$$Q_{ox} \approx (-qt_{ox}N_{t0})[-\xi_n + \xi_n\Gamma(1-a)(Q_{inj}\sigma_0/q)^a] \quad (2.45)$$

Der Formel Ausdruck zeigt, dass die Oxidladung von den Parametern σ_0 , N_{it} und Q_{inj} abhängt und eine Funktion von $(Q_{inj}\sigma_0/q)$ ist. Der Exponent a hängt von zwei oxid-spezifischen Größen L_n und L_σ ab, die die Ortsabhängigkeit der Defekte und den Wirkungsquerschnitt bestimmen.

2.7 Zuverlässigkeitsverhalten von Speicherzellen

Die Zuverlässigkeit von Speicherzellen hängt wesentlich von der Zuverlässigkeit der verwendeten Oxide ab, die durch die Erzeugung von Defektzentren und Oxidladungen begrenzt ist. Daher soll der Einfluss von Defekten und Oxidladungen auf den Datenerhalt und die Datenwechselstabilität von Speicherzellen eingehend erläutert und diskutiert werden.

2.7.1 Datenwechselstabilität

Häufiges Programmieren und Löschen von Speicherzellen schädigt die Tunneloxide der Speicherzellen, wodurch Defekte erzeugt werden und Oxidladungen beider Polarität entstehen können.

In Gegenwart negativer Oxidladungen verringert sich die Spannung über dem Tunneloxid, was die Abnahme des Fowler-Nordheim-Stroms bewirkt. Hierdurch nimmt das Schwellspannungsfenster kontinuierlich ab, bis die Detektionsgrenzen für die Schwellspannungen unterschritten sind und die logischen Zustände nicht mehr unterscheidbar sind (Datenverlust). Durch positive Oxidladungen findet hingegen eine Aufweitung des Schwellspannungsfensters statt, was hauptsächlich im Bereich geringer Oxiddegradationen bzw. Zykelzahlen beobachtet wird. In Bild 2.24 a) ist die typische Cha-

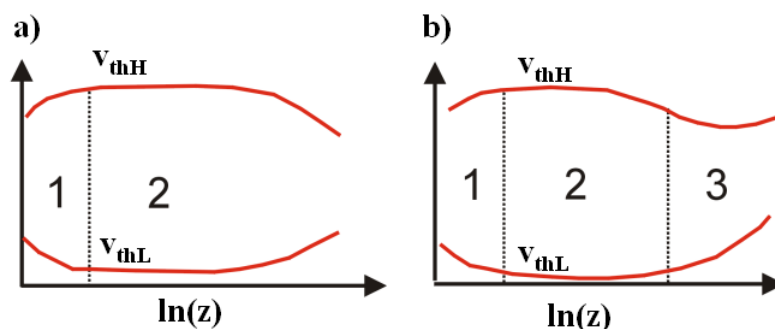


Bild 2.24: a) Schematische Darstellung der Datenwechselstabilität so wie sie bei Single-Polysilizium Speichern und FOTOX-Speicherzellen vorliegt. b) Zeigt die Datenwechselstabilität von Flash-Speicherzellen.

rakteristik einer Datenwechselstabilität skizziert, so wie sie bei FLOTOX oder Single-

Polysilizium-Speicherzellen [69] zu finden ist. Die Schwellspannungen v_{thH} und v_{thL} sind als Funktion der Zykelzahl z dargestellt. Hierbei gibt die Zykelzahl z die Anzahl der Programmier- und Löschyklen an, wobei ein P/E-Zyklus der Abfolge aus Programmieren und Löschen entspricht. Die Programmier- und Löschyklen führen im Bereich geringer Ladungsinjektion (Bereich 1) zu einer Aufweitung des Schwellspannungsfensters, da im Bereich schwacher Degradationen der Einfang von positiven Ladungen vorherrschend [70,71] ist. Im Bereich 2 dominiert der Einfang von negativen Oxidladungen, die zu einer Reduktion des Schwellspannungsfensters führen. Die hierbei entstehenden Charakteristiken können symmetrisch oder asymmetrisch sein, was sowohl von den Bedingungen beim Programmieren und Löschen als auch von der Position des Ladungsschwerpunktes abhängt. Wenn der Ladungsschwerpunkt nicht mittig im Oxid liegt, sind die elektrischen Felder richtungsabhängig, was sich unmittelbar auf die Symmetrie der Schwellspannungsverläufe auswirkt. In Bild 2.24 b) ist die Datenwechselstabilität einer Speicherzelle skizziert, bei der die Oxidladungen sowohl den Fowler-Nordheim-Strom als auch die Schwellspannung des Auslesetransistors beeinflussen. Dieser Fall liegt z. B. bei Flash-Speicherzellen vor, da die Programmierung bzw. Löschung der Speicherzelle über den n-Kanal des Auslesetransistors stattfindet. Der Ladungseinbau erfolgt somit im Gateoxid des Auslesetransistors. In den Bereichen 1 und 2 findet durch positive bzw. negative Oxidladung eine Aufweitung bzw. Reduktion des Schwellspannungsfensters statt. Die negativen Oxidladungen bewirken zusätzlich eine Erhöhung der Schwellspannung des Auslesetransistors v_{thN} , was im dritten Teilabschnitt von Bild 4.24 b) deutlich wird. In diesem Bereich beginnt die Schwellspannung anzusteigen, da deren Abnahmen durch die Zunahme von v_{thN} überkompensiert wird. Ein Beispiel hierfür ist in Bild 2.25 zu finden. Hierzu wurden Programmier-

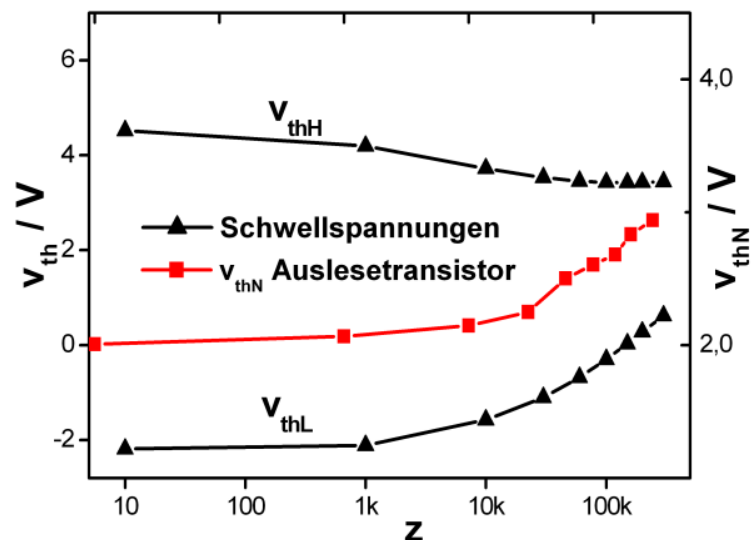


Bild 2.25: Datenwechselstabilität einer Flash-Speicherzelle. Das Schwellspannungsfenster nimmt deutlich mit der Zykelzahl ab, wobei die neutrale Schwelle (rechte Achse) mit der Zykelzahl ansteigt. Ab Zykelzahlen größer 200.000 P/E-Zyklen beginnt die Schwellspannung im programmierten Zustand anzusteigen.

2 Grundlagen zu Speicherzellen

und Löschkzyklen an einer Flash-Speicherzelle ausgeführt und die Schwellspannungen in Abhängigkeit von der Zykelzahl z dargestellt. Deutlich zeigt sich die Abnahme der Schwellspannung im programmierten Zustand mit zunehmender Zykelzahl. Ab 100.000 P/E-Zyklen beginnt die Schwellspannung v_{thH} leicht anzusteigen, was auf den Anstieg der Schwellspannung des Auslesetransistors (rechte Achse) zurückzuführen ist.

2.7.2 Datenerhalt

Der Datenerhalt erfasst die Zeitdauer, mit der eine Information in einer Speicherzelle gespeichert werden kann. Informationsverlust ist auf den Verlust von Floating-Gate-Ladung zurückzuführen, die über das Tunneloxid zwischen Floating-Gate und Substrat abfließen kann, Bild 2.26 a). Eine Entladung des Floating-Gates und eine Änderung der Schwellspannungen bis hin zum Datenverlust sind die Folge, Bild 2.26 b). Wird z. B. eine Speicherzelle mit einer Floating-Gate-Kapazität von 1 pF betrachtet, so dürfen nicht mehr als 2 Elektronen pro Woche verloren gehen, wenn ein Datenerhalt von 10 Jahren erreicht werden soll [72]. Der Ladungsverlust einer Single-Poly-Speicherzelle

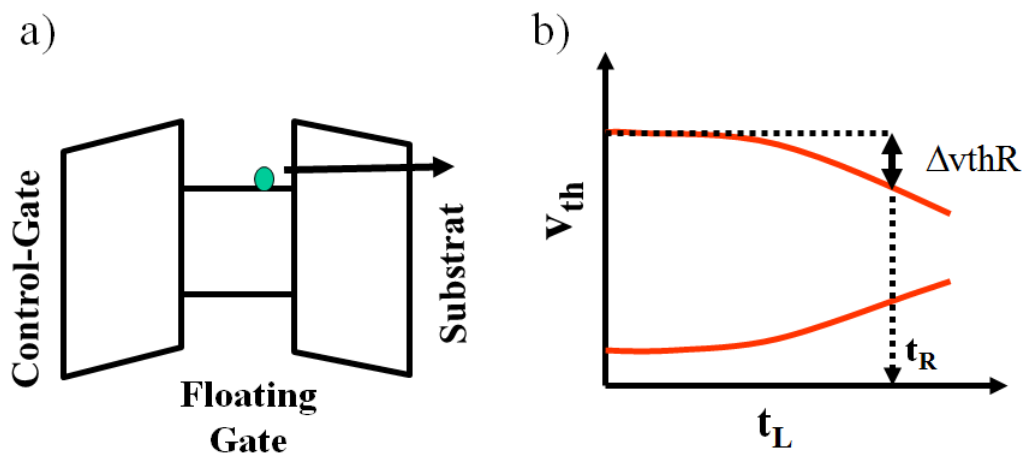


Bild 2.26: a) Schematische Darstellung des Ladungsverlustes einer Speicherzelle im programmierten Zustand. b) Zeigt die Schwellspannungsänderung in Abhängigkeit von der Lagerzeit. Ein Datenverlust liegt vor, wenn die Detektionsgrenzen der Schwellspannungen unterschritten werden.

ist in Bild 2.27 gezeigt. Hierzu wurden programmierte und gelöschte Speicherzellen bei 250 °C gelagert und in definierten Zeitabständen ausgelesen. Die Schwellspannungen sind in Abhängigkeit von der Lagerzeit t_L dargestellt und zeigen ein deutliches Änderungsverhalten, das auf den Verlust von Floating-Gate-Ladung hindeutet. Für die Simulation des Ladungsverlustes müssen die Leckströme bekannt sein bzw. modelliert werden, um hierüber die Schwellspannungen in Abhängigkeit von der Lagerzeit zu simulieren. Bei einem ungealterten Oxid kann die FN-Leitung als Hauptursache für den Datenverlust betrachtet werden. Wenn das Oxid elektrischem Stress ausgesetzt wurde, sind die stressinduzierten Leckströme zu modellieren.

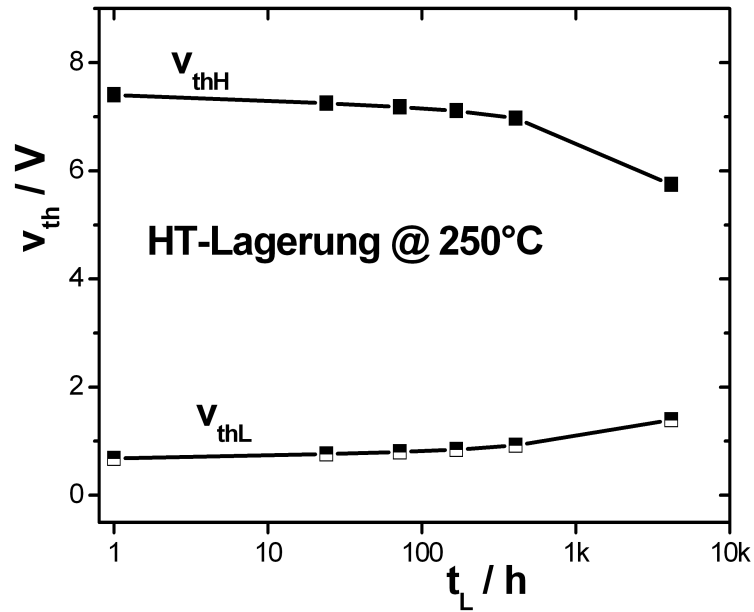


Bild 2.27: Zeigt den Datenverlust einer Single-Poly-Speicherzelle, die bei 250°C gelagert wurde. Die Schwellspannungen im programmierten und gelöschten Zustand sind in Abhängigkeit von der Lagerzeit t_L dargestellt.

Wird ein ungealtertes Oxid betrachtet, so kann über die Fowler-Nordheim-Gleichung die Schwellspannungsänderung in Abhängigkeit von der Lagerzeit bestimmt werden, Gl.2.46 [73]

$$\Delta V_{thR} = \frac{\Delta Q_{FG}}{C_{CG}} = \frac{1}{C_{CG}} \int_0^{t_R} I_{FN}(E_{ox}, t) dt \quad (2.46)$$

Die Zeitdauer t_R wird als Lebensdauer bezeichnet und ΔV_{thR} ist die korrespondierende Änderung der Schwellspannung im programmierten Zustand, siehe Bild 2.26 b). Simulationen mit diesem Modell [73] zeigen, dass der Datenerhalt von verschiedenen Prozessparametern abhängt, wie z. B. der Dotierstoffkonzentration und der Dicke des Tunneloxides. So konnte gezeigt werden, dass mit abnehmender Dotierstoffkonzentration im Injektorgebiet der Datenerhalt verbessert werden kann. Ursache hierfür ist die Abnahme der Fermienergie E_F mit sinkender Dotierstoffkonzentration, womit die Barrierenhöhe zwischen Silizium und Siliziumdioxid steigt und der Datenerhalt verbessert wird. Eine weitere Verbesserung des Datenerhalts kann durch ein dickeres Tunneloxid erreicht werden, da das elektrische Feld über dem Tunneloxid mit zunehmender Oxiddicke abnimmt. Nachteilig ist hierbei, dass mit zunehmender Oxiddicke die Oxiddegradation zunimmt [74-76] und sich die Datenwechselstabilität verschlechtert. Dies zeigt den wechselseitigen Einfluss zwischen Datenerhalt und Datenwechselstabilität, weshalb bei der Optimierung des Datenerhalts meistens auch die Datenwechselstabilität zu berücksichtigen ist.

Bei der experimentellen Bestimmung des Datenerhalts kommen zeitraffende Verfahren zum Einsatz, die durch hohe Temperaturen oder durch elektrische Felder den Ladungsverlust des Floating-Gates beschleunigen. Im Fall hoher Temperaturen werden die Spei-

2 Grundlagen zu Speicherzellen

cherzellen bei unterschiedlichen Temperaturen spannungslos gelagert und in definierten Zeitintervallen ausgelesen. Mithilfe der Arrhenius-Relation [77,78] kann hierüber die Lebensdauer eines Zellzustandes bei Raumtemperatur extrapoliert werden, Gl 2.47 a).

$$a) \quad t_R \propto e^{\frac{\varepsilon_a}{kT}}, \quad b) \quad j = E \sigma_0 \cdot e^{-\varepsilon_a/k_B T} \quad (2.47)$$

wobei ε_a die Aktivierungsenergie, k_B die Boltzmann-Konstante und T die Temperatur ist. Es zeigt sich, dass im Fall kleiner elektrischer Felder und hoher Temperaturen im Siliziumdioxid eine quasi ohmsche Leitung vorliegt [78,79], deren Leitfähigkeit durch eine Arrhenius-Relation beschreibbar ist, siehe Gl.4.47 b). Die elektrische Leitung erfolgt wie beim TAT (Trap-assisted Tunneling) durch die Übergänge von Elektronen zwischen einzelnen Defektniveaus, die im Bereich hoher Temperaturen verstärkt durch thermische Anregung erfolgen. Für die Anwendung der Arrhenius-Relation muss vorausgesetzt werden, dass die Aktivierungsenergie eine temperaturunabhängige Konstante ist. Neuere Untersuchungen haben allerdings ergeben, dass die Aktivierungsenergie im Temperaturbereich oberhalb von 250°C zunimmt, was die Anwendbarkeit der Arrhenius-Relation im Hochtemperaturbereich infrage stellt. Nach De Salvo et al. [79,80] ist hierfür die Temperaturabhängigkeit des Leckstromes verantwortlich, der für hohe Temperaturen näherungsweise exponentiell mit der Temperatur ansteigt. Eine solche Temperaturabhängigkeit impliziert eine Aktivierungsenergie, die quadratisch mit der Temperatur zunimmt Gl.2.48,

$$\varepsilon_a = k_B \frac{T^2}{T_0}, \quad (2.48)$$

wobei T_0 eine feldabhängige Größe ist.

3 Modellierung und Messungen

3.1 Transiente Simulation

In diesem Kapitel soll die Modellierung von Speicherzellen vertieft werden und das Programmier- und Löschverhalten transient simuliert und mit realen Messergebnissen verglichen werden.

Für eine zeitabhängige Simulation des Programmier- und Löschvorgangs müssen die elektrischen Felder, FN-Ströme, die Floating-Gate-Ladung und die Schwellspannungen zeitabhängig berechnet werden. Dies ist in geschlossener Form nicht ohne Vereinfachungen möglich. Zu solchen Vereinfachungen zählt z. B. die Linearisierung des FN-Stromes [81,82] oder die Annahme eines Spannungspulses mit unendlich steilen Flanken [83]. Beide Vereinfachungen haben den Nachteil, dass der Einfluss des zeitabhängigen Spannungssignals auf das Programmieren und Löschen nicht untersucht werden kann. Soll dieser Einfluss berücksichtigt werden, so müssen die Grundgleichungen aus Kapitel 2.3 numerisch gelöst werden. Das Vorgehen hierbei ist im Flussdiagramm in Bild 3.1 dargestellt.

Der Programmier- und Löschvorgang wird in n Zeitschritte unterteilt. Für jeden Zeitpunkt $t(n)$ wird zuerst die Spannung über dem Tunneloxid $V_{ox}(n)$, dann das elektrische Feld und schließlich der FN-Strom bestimmt. Mittels des FN-Stromes werden die Floating-Gate-Ladung und hierüber die Schwellspannungen berechnet.

Über die Floating-Gate-Ladung $Q_{FG}(n+1)$ erfolgt die Berechnung von $V_{ox}(n+1)$, was bereits den $n+1$ -ten Iterationsschritt vorbereitet. Die Inkrementierung der Zeitgröße $t(n)$ schließt den n -ten Iterationsschritt ab. Während des Einschalt- und Ausschaltvorganges wird die Spannung über dem Tunneloxid jeweils mit $\Delta t \cdot a_r$ (Anstiegsflanke) oder $-\Delta t \cdot a_f$ (Abstiegsflanke) inkrementiert.

Für die Bestimmung der Inputdaten müssen die erforderlichen Flächen, Kapazitäten und Kopplungsfaktoren (k_{CG} , k_D) bestimmt werden. Die benötigten Kapazitäten können mit Gl.3.1 über Layoutdaten und Prozessparameter berechnet werden

$$C = \frac{\epsilon_0 \epsilon_{Si} A}{t_{ox}}, \quad (3.1)$$

wobei A für die Fläche der Kapazität steht, t_{ox} die Oxiddicke bezeichnet und ϵ_0 und ϵ_{Si} die Dielektrikumskonstante und die Permittivitätszahl für Silizium sind.

Die Tunnelkonstanten A und B können mit den in Kapitel 2 angegebenen Gleichungen bestimmt werden und sind zusammen mit den anderen Simulationsparametern in Tabelle 3.1 aufgeführt. Für die Berechnung der Tunnelkonstanten wurde für das Verhältnis der effektiven Elektronenmassen ein Wert von $m_{Si} : m_{SiO_2} = 0.5$ angenommen [84,85]. Die Barrierenhöhe zwischen Silizium und Siliziumdioxid wurde $\Phi = 3.14$ eV gesetzt.

3 Modellierung und Messungen

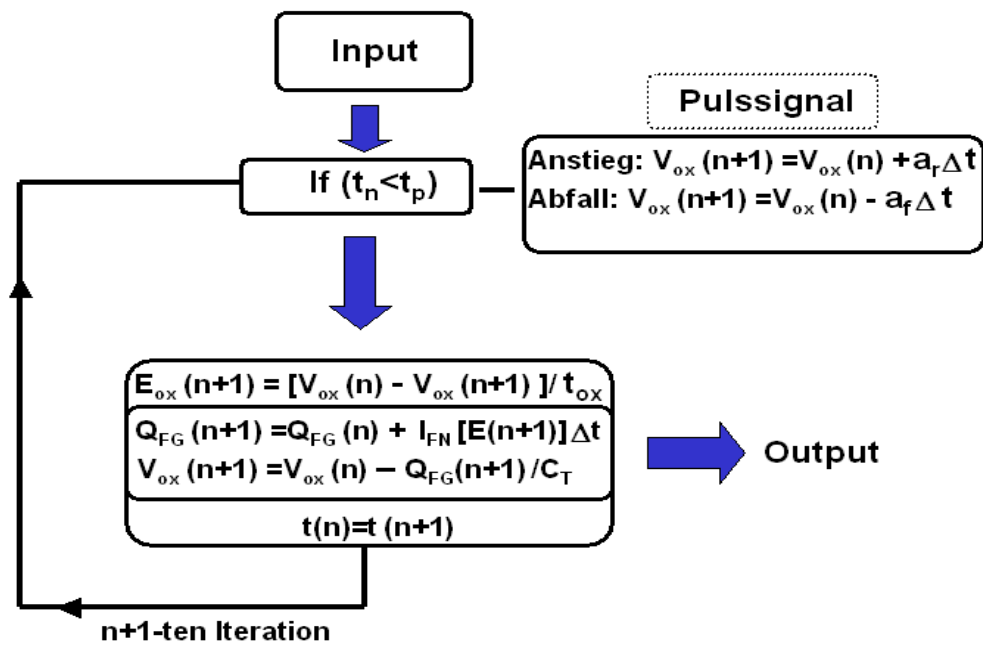


Bild 3.1: Flussdiagramm für die Simulation eines Programmier- und Löschzyklus. Iterativ werden Spannungen, elektrische Felder, Fowler-Nordheim-Ströme und schließlich die Floating-Gate-Ladung berechnet.

	Programmieren	Löschen
FN-Konstanten	$A = 9.98 \cdot 10^{-7} \text{VA}^{-2}$	$A = 9.98 \cdot 10^{-7} \text{VA}^{-2}$
FN-Konstanten	$B = 2.6 \cdot 10^{10} \text{Vm}^{-1}$	$B = 2.6 \cdot 10^{10} \text{Vm}^{-1}$
Kopplung	$k_{CG} = 0.94$	$k_D = 0.026$
Injektorfläche	$A_{inj} = 0.42 \mu\text{m}^2$	$A_{inj} = 0.42 \mu\text{m}^2$
P/E-Spannungen	$V_P = 15\text{V}$	$V^E = 15\text{V}$

Tabelle 3.1: Zusammenstellung der verwendeten Simulationsparameter.

In Bild 3.2 wurden EEPROM-Größen für einen P/E-Zyklus simuliert. Die Zeitdauer des P/E-Zyklus betrug 4 ms und wurde für die folgende Erläuterung in vier Zeitabschnitte unterteilt.

In **Abschnitt 1** beginnt der Programmiervorgang mit dem Anstieg der Control-Gate-Spannung V_{CG} und der Spannung über dem Tunneloxid V_{ox} . Hierzu zeitlich verzögert beginnt der Fowler-Nordheim-Strom zu fließen.

Zu Beginn des **Abschnitts 2** steigt die Schwellspannung des Auslesetransistors an, nachdem der FN-Strom sein Maximum erreicht hat. Gleichzeitig entsteht durch die Floating-Gate-Ladung ein Gegenfeld, das die Spannung über dem Tunneloxid und den FN-Strom abfallen lässt.

In **Abschnitt 3** wird zum Löschen der Speicherzelle am Drainkontakt ein Spannungssignal angelegt. Analog zu Abschnitt 1 steigen V_{ox} und der FN-Strom an. Durch die Umladung des Floating-Gates beginnt die Schwellspannung deutlich zu sinken.

In **Abschnitt 4** führt die Floating-Gate-Ladung erneut zu einem Abfall des Fowler-

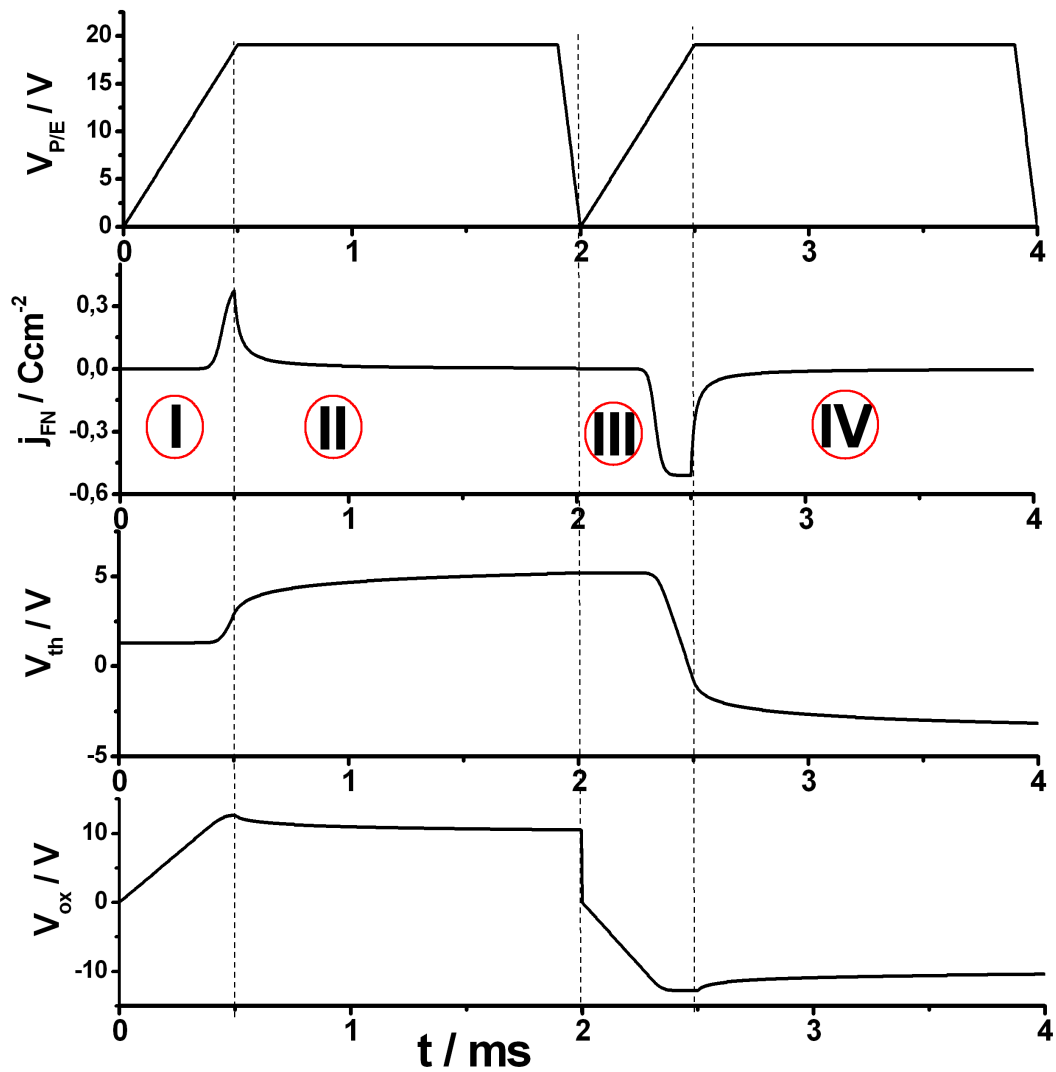


Bild 3.2: Zeitabhängige Simulation von EEPROM-Größen für einen Programmier- und Löschyklus von 4 ms.

Nordheim-Stroms und der Spannung über dem Tunneloxid. In Bild 3.3 ist ein weiteres Simulationsergebnis dargestellt, das die Schwellspannung in Abhängigkeit von der Programmier- und Löschespannung $V_{P/E}$ zeigt. Oberhalb einer P/E -Spannung von 11.5 V liegt eine lineare Abhängigkeit zwischen Schwellspannung und P/E -Spannung vor. Unterhalb von 11.5 V wird ein nichtlineares (exponentielles) Verhalten beobachtet [25], wobei sich die Schwellspannung des programmierten und gelöschten Zustandes an den ladungsfreien Zustand ($v_{thN}=1.3$ V) annähert.

Eine weitere Untersuchung wurde im Hinblick auf die Flankensteilheiten durchgeführt, da diese Größe Einfluss auf die EEPROM-Zuverlässigkeit hat. Hierzu wurden für verschiedene Flankensteilheiten die Stromdichten simuliert und über der Programmier- und Löscheszeit aufgetragen, Bild 3.4. Deutlich ist die Abnahme der Stromdichtemaxima mit abnehmender Anstiegsflanke zu erkennen. Da die Strommaxima am Ende der Anstiegsflanke erreicht werden ($t_r=0.1$ ms, 0.2 ms, 0.3 ms, 0.5 ms), werden die

3 Modellierung und Messungen

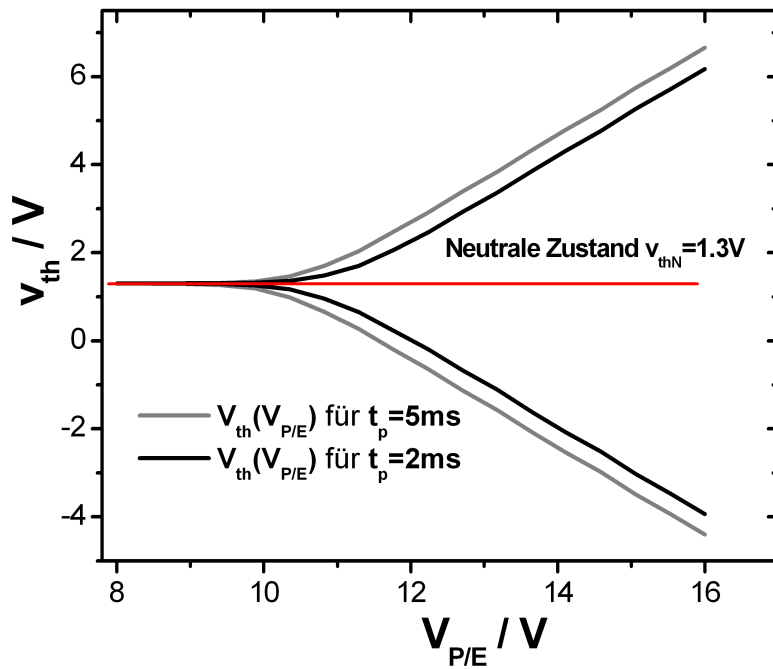


Bild 3.3: Die Schwellspannungen für den programmierten und gelöschten Zustand sind in Abhängigkeit von $V_{P/E}$ simuliert worden. Gezeigt sind zwei Kennlinien für unterschiedliche Pulsängen.

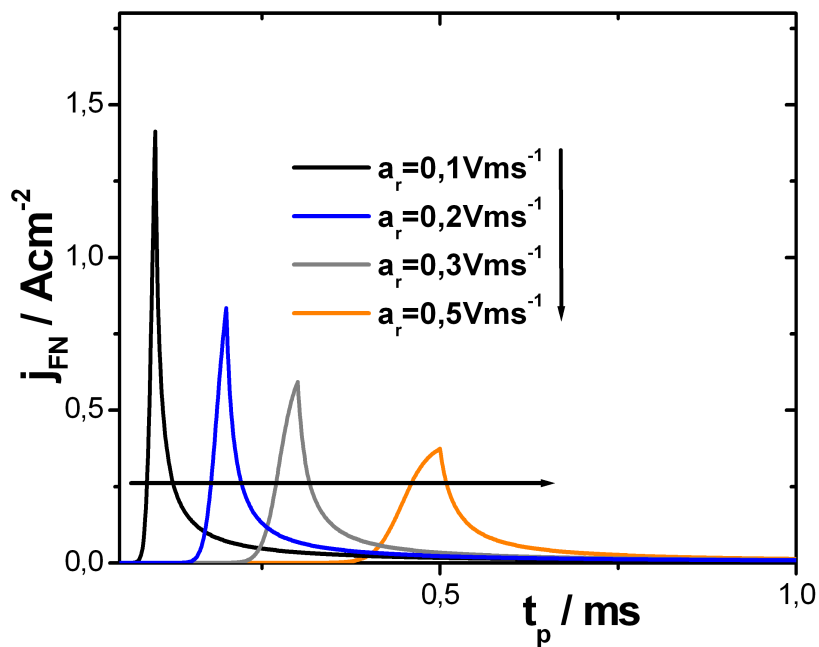


Bild 3.4: Die Simulation zeigt den Fowler-Nordheim-Strom in Abhängigkeit von der Pulszeit t . Dargestellt sind Simulationen für verschiedene Anstiegsflanken.

Strommaxima mit fallender Flankensteilheit nach rechts verschoben. Die Abhängigkeit der Oxiddegradation von der Stromdichte (Stromdichtemaximum) ermöglicht es, die Datenwechselstabilität über die Variation der Flankensteilheit bzw. der Pulsform zu optimieren [81,82].

3.1.1 Simulation mit Oxidladungen

Das entwickelte Simulationsverfahren soll nun unter Berücksichtigung von Oxidladungen erweitert werden. Das in Kapitel 2.3 vorgestellte Modell wird verwendet, um den Einfluss von Oxidladungen auf die Fowler-Nordheim-Injektion zu berücksichtigen. Das Ersatzschaltbild der Speicherzelle sowie des Injektors ist in Bild 3.5 zu sehen. Der Injektor kann durch einen kapazitiven Spannungsteiler beschrieben werden, dessen Enden mit dem Floating-Gate bzw. Drainkontakt verbunden sind. Durch Anwendung des La-

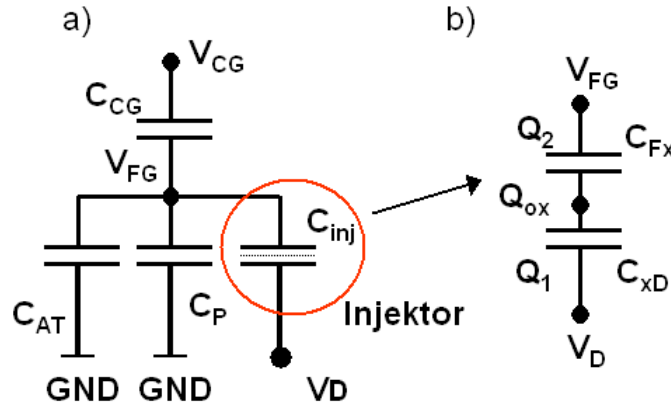


Bild 3.5: a) Zeigt das Ersatzschaltbild der EEPROM-Zelle. An der Injektorkapazität liegt die Spannung $V_{FG} - V_D$ an. b) Zeigt das Ersatzschaltbild des Injektors mit eingebauter Oxidladung.

dungserhaltungssatzes auf den Spannungsteiler in Bild 3.5 b) kann der Spannungsabfall über dem Tunneloxid berechnet werden.

$$V_{ox} - V_D(t) = \frac{Q_{ox}}{C_t} + \frac{C_{Fx}}{C_t} [V_{FG} - V_D(t)] \quad C_t = C_{Fx} + C_{xD} \quad (3.2)$$

Über die Gl. 3.2 lassen sich die elektrischen Felder im Programmier- und Löschfall bestimmen:

$$a) \quad E^E(t) = \frac{V_{ox}^E - V_{FG}}{t_{ox} - x} \quad b) \quad E^P(t) = \frac{V_{ox}^P - V_D(t)}{x}, \quad (3.3)$$

wobei Gl.3.3a) den Programmierfall und Gl.3.3b) den Löschfall beschreibt.

Die elektrischen Felder liefern über die Fowler-Nordheim-Gleichung (Gl.3.4) die Floating-Gate-Ladung.

$$Q_{FG}^{P/E}(t) = \int_0^{t_p} I_{FN}(t) dt = A \int_0^{t_p} (E^{P/E})^2(t) e^{-\frac{B}{E^{P/E}(t) dt}} \quad (3.4)$$

3 Modellierung und Messungen

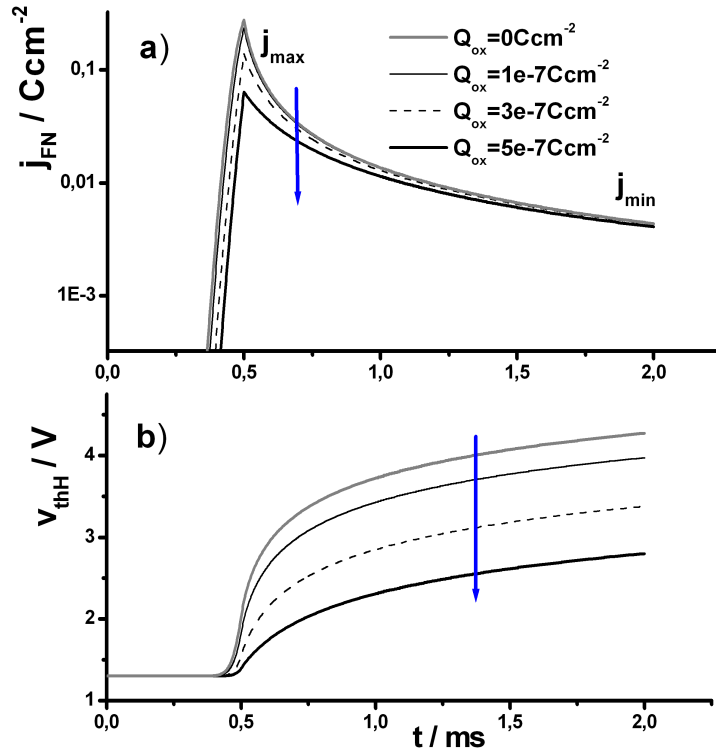


Bild 3.6: a) Simulation der Fowler-Nordheim-Stromdichte für einen Programmierpuls in Abhängigkeit von der Pulszeit t . Es sind Ergebnisse für verschiedene Oxidladungen gezeigt. b) Zeigt die Schwellspannungen des programmierten Zellzustandes in Abhängigkeit von der Pulszeit für verschiedene Oxidladungen.

In Verbindung mit Gl.3.5 [42] ergibt sich hieraus das Floating-Gate-Potenzial,

$$V_{FG}^{P/E} = k_D V_D(t) + k_{CG} V_{CG}(t) + \frac{Q_{FG}^{P/E}(t)}{C_T} + \frac{x \cdot k_D}{A_{inj} \epsilon_0 \epsilon_r} Q_{ox} \quad (3.5)$$

mit der Oxidladung Q_{ox} und dem Ladungsschwerpunkt x . Die Gleichung 3.5 geht für den oxidladungsfreien Fall für $V_D = 0$ in die aus Kapitel 2 bekannte Formel über:

$$V_{FG}^P = k_{CG} V_{CG}(t) + \frac{Q_{FG}(t)}{C_T}$$

Für eine transiente Simulation wurden diese Gleichungen numerisch unter Berücksichtigung des Pulssignals gelöst, wobei die Parameter aus Tabelle 3.1 verwendet wurden. In Bild 3.6a) ist die Zeitabhängigkeit der FN-Stromdichte zu sehen. In Bild 3.6 b) sind die korrespondierenden Schwellspannungen des programmierten Zustandes für verschiedene Oxidladungen dargestellt. Mit zunehmender Oxidladung kann eine deutliche Abnahme der Strommaxima festgestellt werden. Der minimale Stromwert erweist sich hingegen als quasi unabhängig von der eingebauten Oxidladung.

Diese Eigenschaft des Fowler-Nordheim-Stromes überträgt sich auch auf das elektrische Feld über dem Tunneloxid [114,115], sodass auch hier die Oxidladung nur die maximale Feldstärke E_{max} verändert. Wie in Kapitel 4 gezeigt wird, hängen die Floating-

Gate-Ladung und die Schwellspannungen von der minimalen und maximalen Feldstärke ab. Wird nun die Schwellspannung durch die eingebaute Oxidladung verändert, so folgt mit dem Ergebnis aus Bild 3.6, dass die Änderung der Schwellspannung proportional zur Änderung des maximalen elektrischen Feldes ist.

$$v_{th} \propto (E_{max} - E_{min}) \Rightarrow \Delta v_{th}(Q_{ox}) \propto \Delta E_{max}(Q_{ox})$$

3.1.2 Modelloptimierung

Das in den vorangegangenen Abschnitten entwickelte Modell soll anschließend an die hier verwendeten Single-Poly-Speicherzellen angepasst werden.

Neben Layout- und Prozessparametern müssen hierzu weitere Effekte berücksichtigt werden. Beim Löschen der Zelle ist zusätzlich ein Spannungsabfall über den Select-Transistor ΔV_{SG} zu berücksichtigen, siehe Bild 3.7. Da sich im LösCHFall eine Verarmungszone unterhalb des Injektors ausbildet, findet ein weiterer Spannungsverlust statt, der mit ΔV_{dpl} bezeichnet wird.

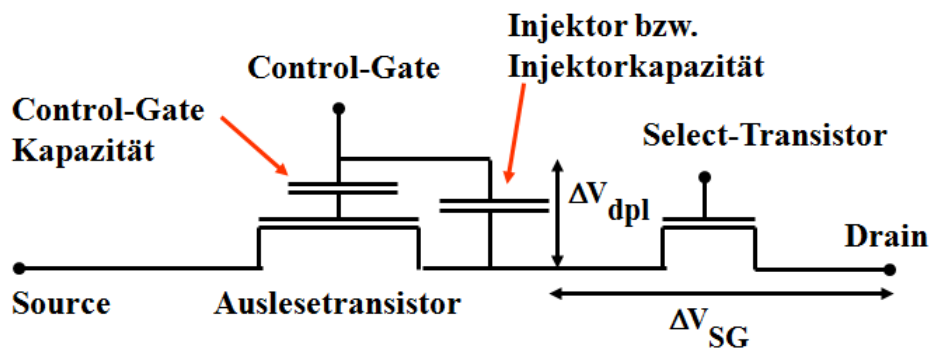


Bild 3.7: Schematische Darstellung der Single-Poly-Speicherzelle. Beim Löschen findet ein Spannungsabfall über dem Select-Transistor ΔV_{SG} statt. Der Spannungsabfall über dem Injektor wird mit ΔV_{dpl} bezeichnet.

3 Modellierung und Messungen

Der Spannungsabfall über dem n-Kanal des Select-Transistors kann anhand eines einzelnen Select-Transistors bestimmt werden. Hierzu wird am Select-Gate und am Drainanschluss die volle Löschspannung angelegt. Wird dem Sourcekontakt ein Drainstrom eingeprägt, so kann durch Messung der Sourcespannung direkt der Spannungsabfall bestimmt werden, den der Kanalstrom bewirkt. Auf diese Weise wurde ein Spannungsabfall von $\Delta V_{SG}=1.7$ V gemessen [86].

Um den Spannungsabfall über der Verarmungszone zu bestimmen, wurden an EEPROM-Injektoren die gate- und substratseitige Fowler-Nordheim-Injektion spannungsabhängig gemessen. Anhand der Spannungsverschiebung der Kennlinien [25] konnte ein Spannungsabfall von $\Delta V_{dpl}=0.3$ V gefunden werden

Das hier entwickelte Modell lässt sich unter Beachtung der effektiven Löschspannung über die Tunnelkonstanten an die gemessenen Schwellspannungswerte anpassen. Hierzu wurden die Schwellspannungen in Abhängigkeit von der P/E-Spannung simuliert und die Fowler-Nordheim-Konstanten B_E und B_P als Fitparameter benutzt. Die Abhängigkeit zwischen den Schwellspannungen und den B-Parametern wird durch einen linearen Zusammenhang beschrieben, was deutlich aus Bild 3.8 hervorgeht. Als Startwerte wurden die folgenden Werte aus [87] benutzt:

Programmierfall: $A = 1.82 \cdot 10^{-7} AV^{-2}$ $B = 1.88 \cdot 10^{10} mV^{-1}$

Löschfall: $A = 1.23 \cdot 10^{-6} AV^{-2}$ $B = 2.37 \cdot 10^{10} mV^{-1}$

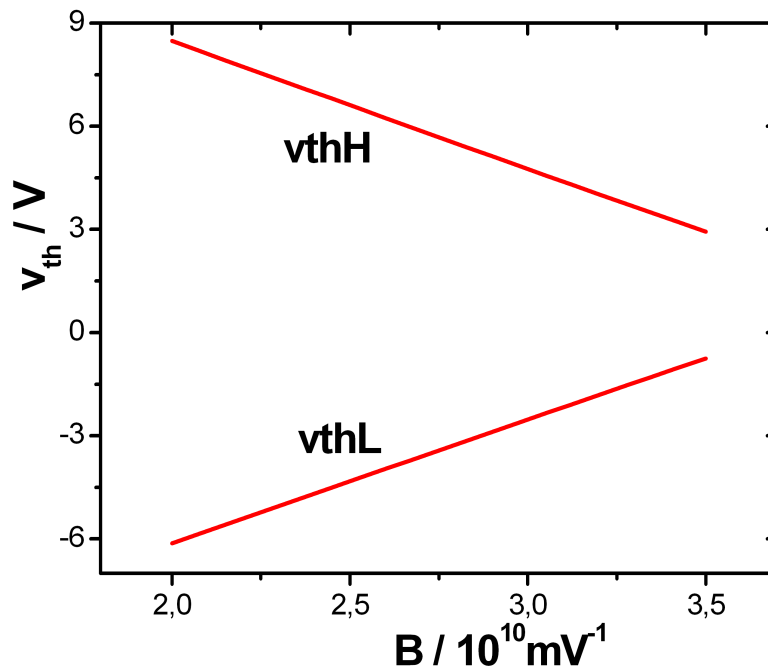


Bild 3.8: Die Schwellspannungen des programmierten und gelöschten Zustandes wurden in Abhängigkeit vom B-Parameter simuliert.

In Bild 3.9 wurde die Schwellspannung als Funktion der P/E-Spannung simuliert. Der B-Parameter wurde hierzu von $2.7 \cdot 10^{10} Vm^{-1}$ bis $3.3 \cdot 10^{10} Vm^{-1}$ variiert wohingegen der A-Parameter konstant gehalten wurde.

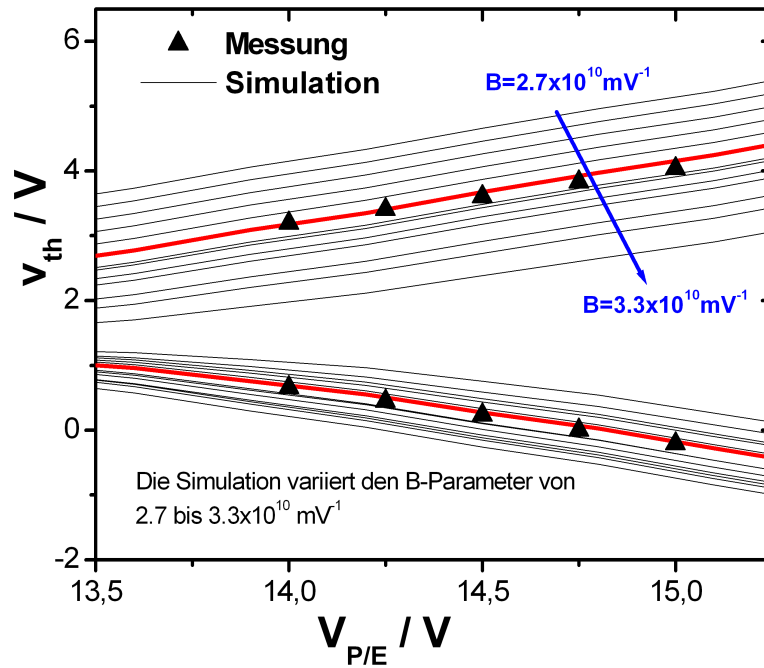


Bild 3.9: Gemessene und für verschiedene B-Parameter simulierte Schwellspannungsverläufe in Abhängigkeit von der P/E-Spannung. Der A-Parameter wurde konstant gehalten.

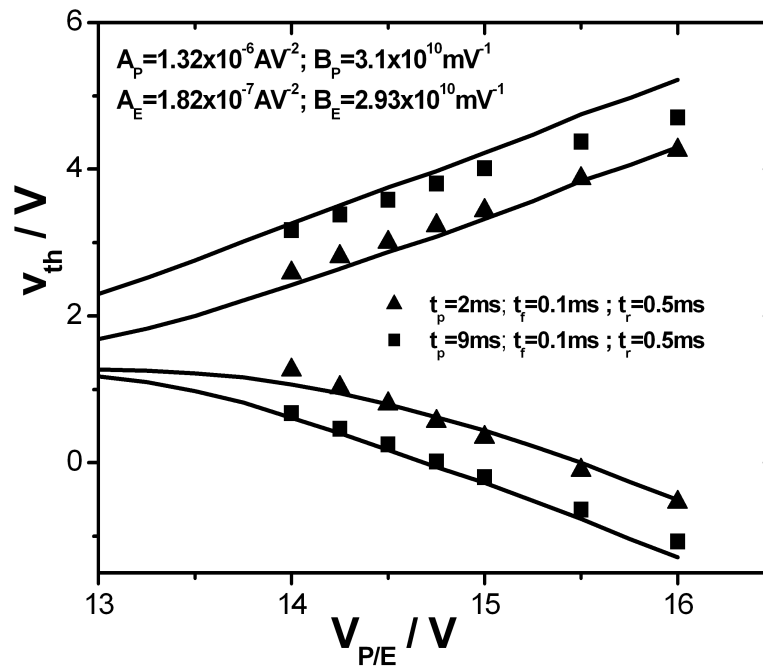


Bild 3.10: Die Graphik zeigt die Schwellspannungen als Funktion der P/E-Spannung. Mit den extrahierten Fitparametern konnte eine gute Übereinstimmung zwischen Messung und Simulation erreicht werden.

3 Modellierung und Messungen

Für den Datensatz:

A-Konstanten: $A_P = 1.82 \cdot 10^{-7} \text{VA}^{-2}$ und $A_E = 1.23 \cdot 10^{-6} \text{VA}^{-2}$

B-Konstanten: $B_P = 2.98 \cdot 10^{10} \text{Vm}^{-1}$ und $B_E = 3.1 \cdot 10^{10} \text{Vm}^{-1}$

ließ sich eine gute Übereinstimmung zwischen Messung und Simulation erreichen, was Bild 3.9 zeigt.

Weitere Messungen an EEPROM, die Datensätze zu zwei unterschiedlichen Pulslängen zeigen, sind in Bild 3.10 dargestellt. Es zeigt sich, dass die Abhängigkeit der Schwellspannungen von der P/E -Spannung sowie von der Pulslänge richtig beschrieben wird und die extrahierten Fitparameter die Messdaten mit hinreichender Genauigkeit modellieren. Lediglich bei der Pulslänge von 9 ms weicht die Simulation für 15.5 V und 16 V geringfügig von den Messwerten ab.

3.2 Messungen

3.2.1 Aufbau der Messplätze

Nach den Ausführungen zur Modellierung von Speicherzellen werden Messungen zum Programmier- und Löschverhalten sowie zur Datenwechselstabilität von Single-Poly-Speicherzellen gezeigt und diskutiert. Die für die Messungen verwendeten Messplätze sollen zuvor erläutert werden.

Für die Messungen an Speicherzellen standen ein NVM-Messplatz und ein Spitzenmessplatz zur Verfügung. Der NVM-Messplatz wurde für Messungen an 384Bit-Arraystrukturen verwendet und ermöglichte die automatisierten Messungen von Wafern und Chargen. Am Spitzenmessplatz fanden exemplarische Messungen statt, die hohe Temperaturen bis 300°C und hohe Messgenauigkeit erforderten.

Der NVM-Messplatz ist schematisch in Bild 3.11 dargestellt. Er besteht aus einem PC, einem Pulsgenerator, einem HP-Analyser, einer Matrixverschaltung und einem voll automatisierten Messautomaten. Sämtliche Messgeräte sind über einen GPIB-Bus miteinander verbunden und können über einen PC angesprochen werden. Zu den Messgeräten gehört ein Pulsgenerator (TTi-Abitrary Waveform Generator: TGA 1240), mit dem sich beliebig geformte Spannungspulse generieren lassen. Für die Programmierung und Löschung der Single-Poly-Speicherzelle wurden trapezförmige Spannungspulse erzeugt, die durch eine Anstiegsflanke, Plateaubereich und eine Abtriegsflanke charakterisiert sind und deren Pulslängen im Bereich einiger Millisekunden liegen. Als Standardpuls für die Programmierung und Löschung wird ein Spannungssignal mit $t_r = 0.5 \text{ms}$, $t_f = 0.1 \text{ms}$ und variablem Plateaubereich bzw. Spannungsamplitude definiert. Die Ausgangssignale für Drain, Select-Gate und Control-Gate müssen, bevor sie an die Matrixverschaltung gelangen, verstärkt werden, da die Ausgangsspannung des Pulsgenerators auf 10 V begrenzt ist und für die Programmierung und Löschung der Speicherzellen deutlich höhere Spannung erforderlich ist. Der HP-Analyser (HP4142) wird für die Messung von Spannungen und Strömen verwendet und verfügt über vier Source-Measurement-Units (SMU1, SMU2, SMU3, SMU4) über die Ströme und Spannungen eingepreßt und gemessen werden können. Zusätzlich stehen eine Ground (GND)-Unit, Hochvolt-Unit und eine AFU (Analog-Feedback-Unit) zur Verfügung. Über die AFU können die Messungen der EEPROM-Schwellspannungen deutlich schneller durchge-

führt werden als über eine getrennte Steuerung der SMU-Einheiten des Analysers. Hierzu wird einem Force-Kanal eine Spannung eingepreßt und der Sens-Wert einer anderen SMU wird mit einem Referenzwert verglichen, bis dieser erreicht ist. Die so ermittelte Spannung wird als Schwellspannungswert ausgegeben. Über die Matrixverschaltung (Switch-Mainframe) werden die SMUs des HP-Analysers und die Eingänge des Puls-generators mit der Nadelkarte verschaltet, die auf dem EEPROM-Array aufsetzt. Das Switch-Mainframe verfügt über 10 Eingänge, von denen 8 Eingänge belegt sind (4 SMUs, 3 Kanäle für den Puls-generator und ein GND-Anschluss). So ist gewährleistet, dass jede SMU und jeder Kanal des Puls-generators mit jeder Kontaktfläche der Array-struktur verbunden werden kann.

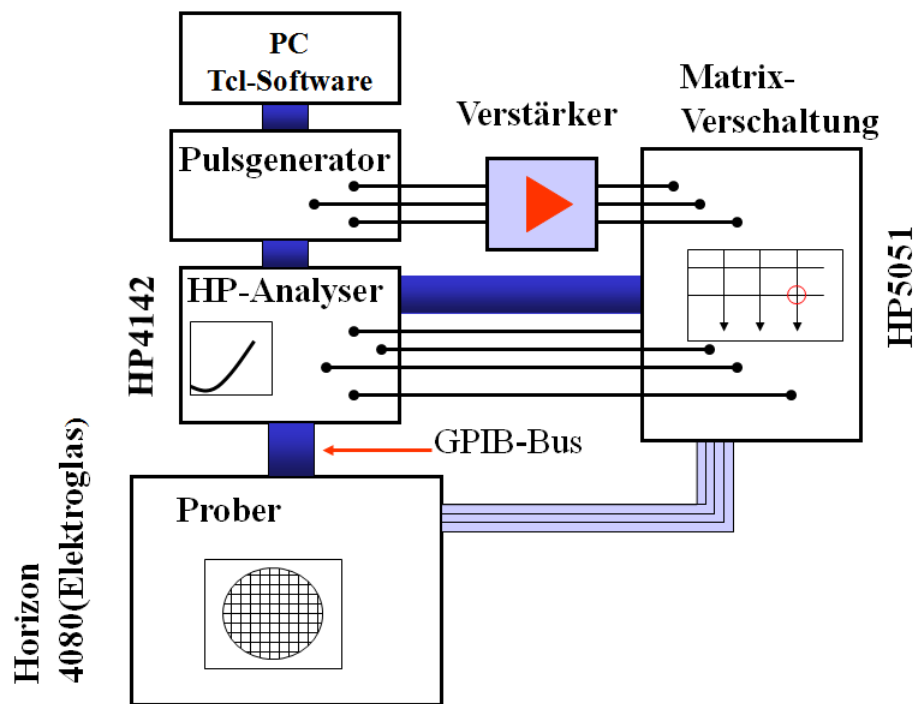


Bild 3.11: Die Abbildung zeigt den schematischen Aufbau des verwendeten NVM-Messplatzes. Die Messgeräte sind über einen GPIB miteinander verbunden und sind über einen PC ansteuerbar. Der Messautomat (Prober) ermöglicht die automatisierte Messung von Wafern und Chagen.

Der Spitzenmessplatz diente zur Charakterisierung von EEPROM-relevanten Bauelementen und einzelnen Speicherzellen, die nicht in Arraystrukturen integriert waren. Der Messplatz befindet sich in einer verschließbaren Box und besteht aus einem HP-Analyser (HP4145), einem Lichtmikroskop und einem beheizbaren Probesteller, der sich zur schwingungsarmen Messungen auf einer Granitplatte befindet.

3 Modellierung und Messungen

3.2.2 EEPROM-Teststruktur

Für die Untersuchung des Programmier- und Löschverhaltens wurden die Speicherzellen eines 384Bit-Arrays unter verschiedenen Konditionen programmiert und gelöscht. Die Teststruktur für die Single-Poly-Speicherzelle ist schematisch in Bild 3.12 dargestellt und besteht aus paarweise gespiegelten Single-Poly-Speicherzellen. Jede Zeile des Arrays kann über einen Control-Gate-Anschluss adressiert werden, wobei die Speicherzellen innerhalb der Zeile über einen Select-Transistor und einen Drainanschluss ansteuerbar sind. Die Speicherzellen einer Wortleitung teilen sich einen Control-Gate-

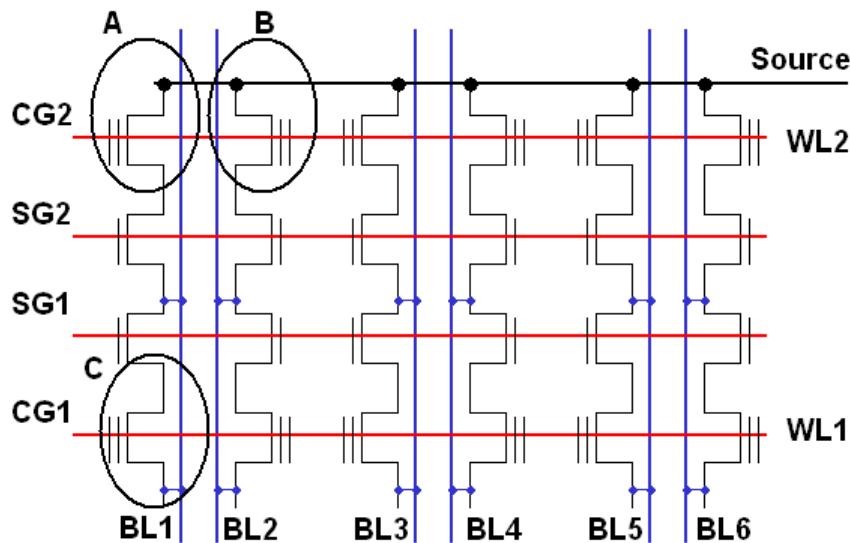


Bild 3.12: EEPROM-Teststruktur des verwendeten 384Bit-Arrays. Es besteht aus paarweise gespiegelten Speicherzellen. Die Drainanschlüsse der Select-Transistoren bilden die Bitleitung (BL), die Control-Gate Anschlüsse der Speicherzelle die Wortleitung (WL).

Anschluss. Die Bitleitung besteht aus den Drainanschlüssen der Select-Transistoren. Insgesamt können über 8 Control-Gates, 6 Select-Gates und 8 Drainanschlüsse 384 Bit angesteuert werden.

Für die Bestimmung der Schwellspannungen werden die Speicherzellen, wie in Tabelle 3.2 für den Lesefall beschrieben, selektiert und beschaltet. Hierzu wird mithilfe der AFU-Einheit des HP-Analysers die Control-Gate-Spannung so lange variiert, bis ein Drainstrom von $5 \mu\text{A}$ erreicht ist. Der Reihe nach werden dann die Drainkontakte durchgeschaltet und die Schwellspannungen der kompletten Zeile ermittelt. Die so bestimmten Schwellspannungen werden für die Definition folgender statistischer Größen herangezogen:

Mittelwert:

$$a) \quad v_{thH} = \frac{1}{384} \sum_{n=1}^{384} v_{thH}(n) \quad b) \quad v_{thL} = \frac{1}{384} \sum_{n=1}^{384} v_{thL}(n) \quad (3.6)$$

Modus	CG	SG	Drain	Source	Bulk
Prog.(A)	V_{pp}	V_{pp}	0V	float.	0V
Löschen(A)	0V	V_{pp}	V_{pp}	float.	0V
Löschen(B)	0V	V_{pp}	0V	float.	0V
Löschen(C)	0V	0	V_{pp}	float.	0V
Read(A)	V_R	V_{DD}	0.5V	0V	0V
Read(B)	V_R	V_{DD}	0V	0V	0V
Read(C)	0V	0V	0.5V	0V	0V

Tabelle 3.2: Zeigt die Signalbelegung der Single-Poly-Speicherzellen für den Programmier- und Löschofall. Zusätzlich sind die Lese-Konditionen aufgeführt. Neben der selektierten Zelle (A) sind bei den Löscho- und Lesekonditionen auch die benachbarten nicht selektierten Zellen B und C aufgeführt, siehe Bild 3.12.

Weitere statistische Größen sind das Minimum und Maximum des 384Bit-Arrays ($v_{th,min}$ und $v_{th,max}$) sowie die Standardabweichungen der Schwellspannungen S_{thH} und S_{thL} :

$$a) \quad S_{thH} = \sqrt{\frac{1}{384} \sum_{n=1}^{384} (v_{thH}(n) - v_{thH})^2} \quad b) \quad S_{thL} = \sqrt{\frac{1}{384} \sum_{n=1}^{384} (v_{thL}(n) - v_{thL})^2} \quad (3.7)$$

3.2.3 Programmier- und Löschoverhalten

Nachfolgend wird das Programmier- und Löschoverhalten von Speicherzellen untersucht. Hierzu wurden die Schwellspannungswerte eines 384Bit-Arrays gemessen und die berechneten Mittelwerte als Funktion der P/E-Spannung oder der Pulslänge dargestellt.

In Bild 3.13 sind die Mittelwerte der Schwellspannungen als Funktion der Programmier- und Löscho-Spannung dargestellt. Die lineare Abhängigkeit zwischen Schwellspannung und P/E-Spannung ist deutlich zu erkennen [25,88]. Der lineare Zusammenhang zwischen beiden Größen folgt unmittelbar aus dem Kondensatorgesetz ($V_{P/E} \propto Q_{FG}$) und dem linearen Zusammenhang zwischen Floating-Gate-Ladung und Schwellspannung. Die Messdaten zeigen, dass eine Änderung der Programmier- und Löscho-Spannung von 1 V eine Schwellspannungsänderung von ca. 0.9 V hervorruft. Dieser Befund ist konsistent mit den Simulationsergebnissen aus Kapitel 3.1 (Bild 3.3), die eine Spannungsänderung von 1 V liefern.

Die Abhängigkeit zwischen Schwellspannung und Pulslänge ist in Bild 3.14 dargestellt, wobei die P/E-Spannung konstant bei 15 V liegt. Die Resultate zeigen einen linearen Zusammenhang zwischen dem Logarithmus der Pulslänge und den Schwellspannungen des programmierten und gelöschten Zustands. Eine genaue Betrachtung der Schwellspannungen zeigt, dass nach 9 ms eine Schwellspannung von $v_{thH}=3.65$ V erreicht ist. Hiervon sind bereits 2.67 V nach 1 ms erreicht, was auf die hohe zeitliche Änderung des Fowler-Nordheim-Stromes für Zeiten ≤ 1 ms hindeutet. Dieses Verhalten geht auch aus Arbeiten zu zeitabhängigen Messungen von Tunnelströmen hervor [89,133], die

3 Modellierung und Messungen

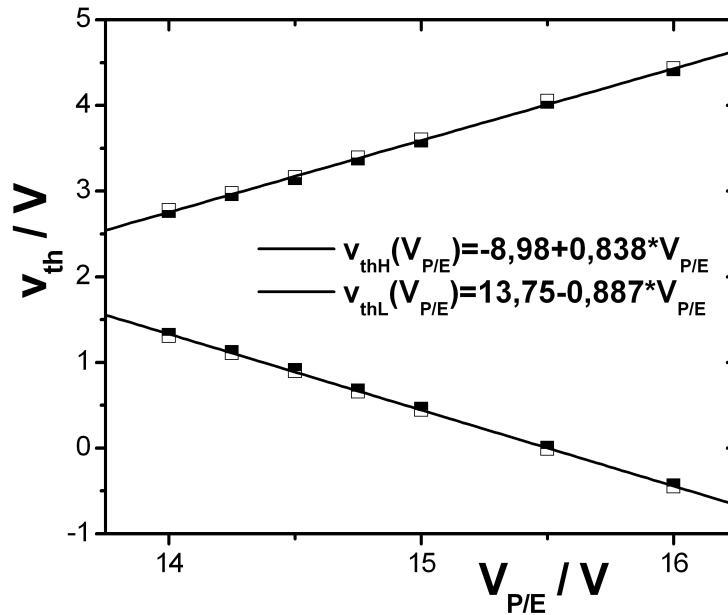


Bild 3.13: Die Schwellspannungen wurden in Abhängigkeit von $V_{P/E}$ gemessen. Deutlich ist der lineare Zusammenhang zwischen Schwellspannung und P/E-Spannung erkennbar.

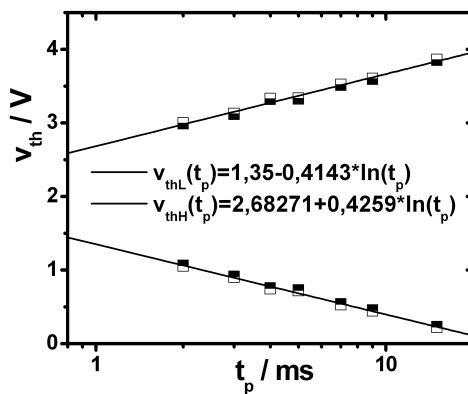


Bild 3.14: Die Schwellspannungen hängen linear von $\ln(t_p)$ ab. Die P/E-Spannung betrug 15 V und die Anstiegs- und Abstiegszeiten entsprachen den Standardeinstellungen.

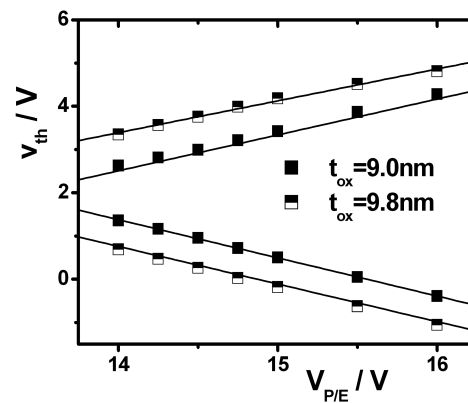


Bild 3.15: Für unterschiedliche Oxiddicken wurden die Schwellspannungen in Abhängigkeit von der P/E-Spannung gemessen.

zeigen, dass die meiste Floating-Gate-Ladung innerhalb der ersten Millisekunde fließt. Weitere Untersuchungen betrachten den Einfluss unterschiedlicher Oxiddicken und Injektorflächen auf den Zusammenhang zwischen Schwellspannung und der P/E-Spannung. In Bild 3.15 wurden an Speicherzellen mit unterschiedlichen Oxiddicken die Schwellspannungen in Abhängigkeit von $V_{P/E}$ gemessen. Die Messungen zeigen eine merkliche Reduktion des Schwellspannungsfensters mit zunehmender Oxiddicke, die durch die Abnahme des Fowler-Nordheim-Stromes verursacht wird. Die Vergrößerung der Injektorfläche führt ebenfalls zu einer Abnahme des Schwellspannungsfensters, siehe Bild 3.16. An Speicherzellen mit unterschiedlich großen Injektorflächen wurde das Schwellspannungsfenster als Funktion der P/E-Spannung gemessen und eine Abnahme des Schwellspannungsfensters mit der Vergrößerung der Injektorfläche festgestellt. Die Reduktion des Schwellspannungsfensters ist auf die Abnahme des Control-Gate-Kopplungsfaktors zurückzuführen, da durch die Vergrößerung der Injektorfläche die Gesamtkapazität steigt und der Control-Gate-Kopplungsfaktor sinkt. Durch Abnahme des Kopplungsfaktors nimmt das Schwellspannungsfenster ebenfalls ab, d. h., es verkleinert sich mit der Vergrößerung der Injektorfläche.

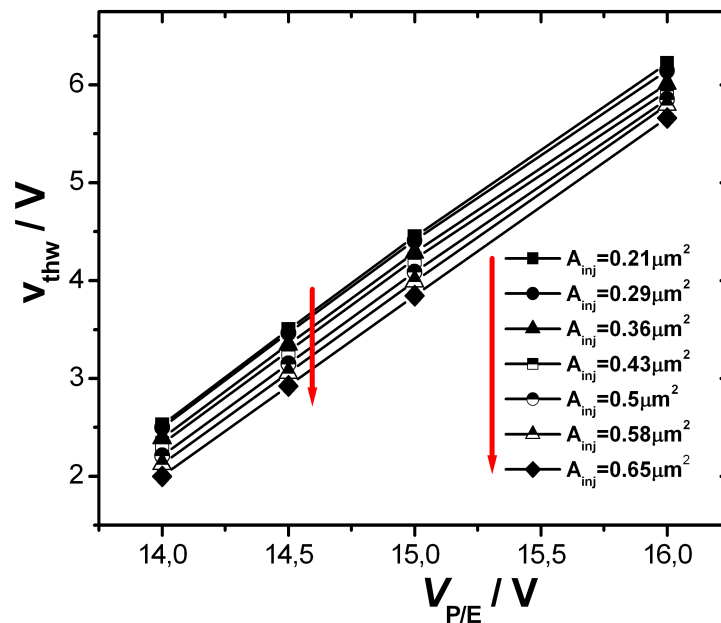


Bild 3.16: Zeigt das Schwellspannungsfenster von Speicherzellen mit unterschiedlichen Injektorflächen: Das Schwellspannungsfenster wurde in Abhängigkeit von der P/E-Spannung dargestellt.

3.2.4 Messungen zur Datenwechselstabilität

Im Folgenden soll der Einfluss von Oxidladungen auf die Schwellspannungen untersucht werden. An Speicherzellen wurden Programmier- und Löszyklen durchgeführt und nach vordefinierten Zykelzahlen unterbrochen, um Schwellspannungen und Transferkennlinien des Auslesetransistors zu bestimmen. In Bild 3.17 sind die Transferkennlinien des Auslesetransistors für verschiedene Zykelzahlen dargestellt. Mit zunehmenden

3 Modellierung und Messungen

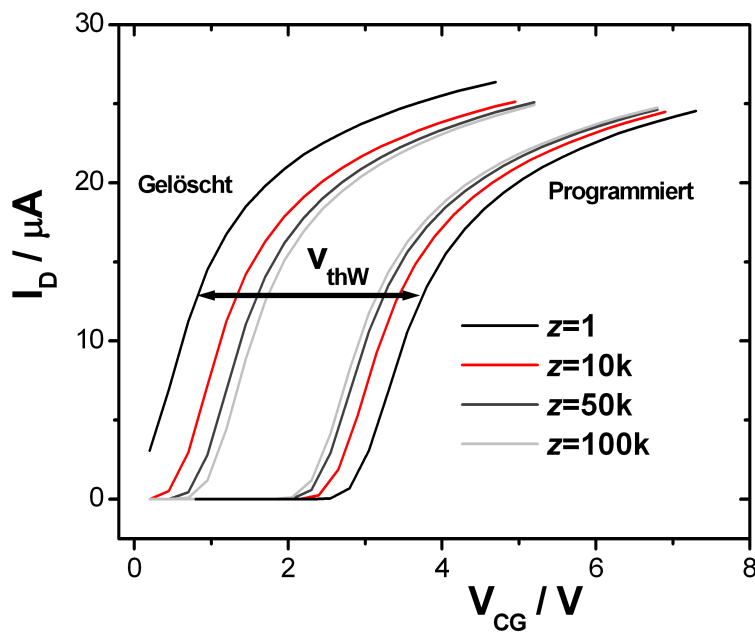


Bild 3.17: Der Abstand zwischen den Transferkennlinien des Auslesetransistors nimmt mit steigender Zykelzahl ab. Dies entspricht der Abnahme des Schwellspannungsfensters v_{thW} mit steigender Zykelzahl.

der Zykelanzahl verkleinert sich das Schwellspannungsfenster, das dem Abstand zwischen den Transferkennlinien entspricht. Die Abnahme des Schwellspannungsfensters geht noch deutlicher aus dem Verhalten der Schwellspannung als Funktion der Zykelzahl hervor, Bild 3.18. Der Einbau negativer Oxidladungen verringert merklich das Schwellspannungsfenster, wobei die Schwellspannung im programmierten Zustand eine negative Änderung ($\Delta v_{thH} < 0$) und im gelöschten Zustand eine positive Änderung erfährt $\Delta v_{thL} > 0$.

Neben dem Einbau von negativen konnte auch der Einfluss von positiven Oxidladungen festgestellt werden. Die Bildung von positiven Oxidladungen findet hauptsächlich im Bereich schwacher Degradation statt und kann auf den Einfang von Löchern oder auf die Entleerung (Detrapping) von Defektniveaus zurückgeführt werden [90,91]. Das Teilbild in Bild 3.18 zeigt den Schwellspannungsverlauf des programmierten Zustandes für $z \leq 300$. Eine deutlich erkennbare Aufweitung des Schwellspannungsfensters kann beobachtet werden, die auf positive Oxidladungen hindeutet. Bei den hier verwendeten Oxiden von 9.4 nm befindet sich die Aufweitung des Schwellspannungsfensters in der Größenordnung von etwa 50 mV und ist bereits nach 10 bis 20 P/E-Zyklen erreicht.

Die Abhängigkeit der Datenwechselstabilität von der Programmier- und Löschspannung zeigt Bild 3.19. Um den Einfluss der P/E-Spannung auf die Oxiddegradation erkennen zu können, wurden die Spannungsänderungen Δv_{thH} und Δv_{thL} als Funktion der Zykelzahl dargestellt. Die Ergebnisse zeigen eine deutliche Zunahme der Schwellspannungsänderung bzw. der Oxiddegradation mit steigender Spannung.

In Bild 3.20 ist die Datenwechselstabilität von Speicherzellen mit unterschiedlich großen Injektorflächen zu sehen. Um eine bessere Vergleichbarkeit der Messergebnisse zu ermöglichen, wurden die Speicherzellen so programmiert und gelöscht, dass die erreich-

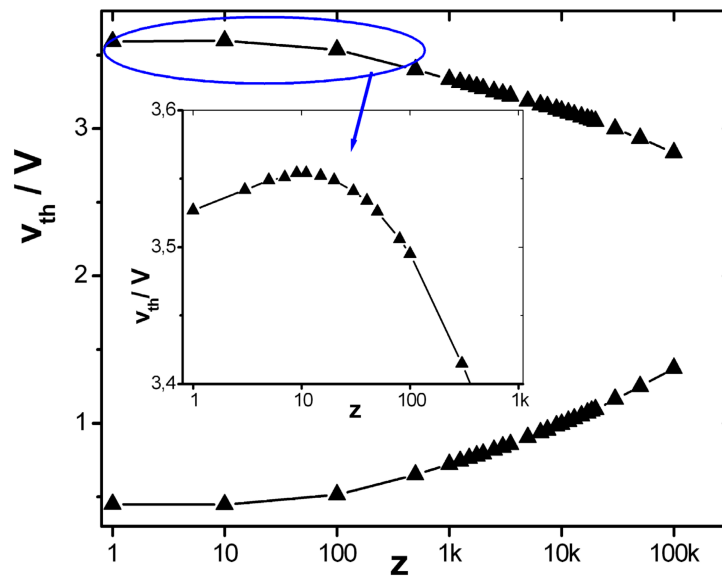


Bild 3.18: Zeigt das Änderungsverhalten der Schwellspannungen in Abhängigkeit von der Zykelzahl. Im programmierten (gelöschten) Zustand kann eine negative (positive) Änderung der Schwellspannungen festgestellt werden. Im Bereich der schwachen Degradation (Teilbild) wird die Schwellspannung im programmierten Zustand geringfügig angehoben.

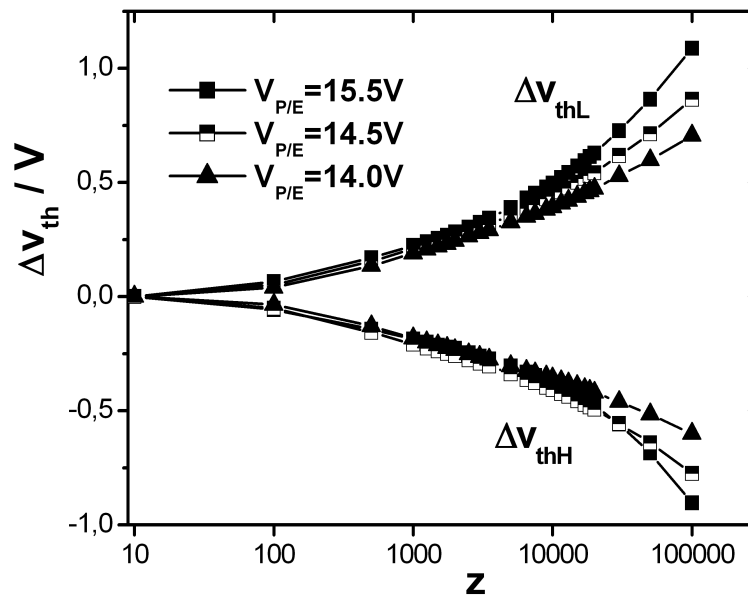


Bild 3.19: Es wurde der Einfluss von unterschiedlichen Programmier- und Löschspannungen auf das Änderungsverhalten der Schwellspannungen untersucht. Die Schwellspannungsänderungen sind über der Zykelzahl aufgetragen und nehmen deutlich mit steigender P/E-Spannung zu.

3 Modellierung und Messungen

ten Schwellspannungsfenster gleich waren; deutlich zeigt sich, dass mit zunehmender Injektorfläche die Oxiddegradation steigt und sich das Schwellspannungsfenster schneller schließt. Mit der Vergrößerung der Injektorfläche nimmt der Control-Gate-Kopplungsfaktor ab, sodass die P/E-Spannung erhöht werden muss, um ein spezifiziertes Schwellspannungsfenster zu erreichen. Aufgrund der höheren P/E-Spannung nimmt somit die Oxiddegradation mit der Vergrößerung des Injektorfensters zu.

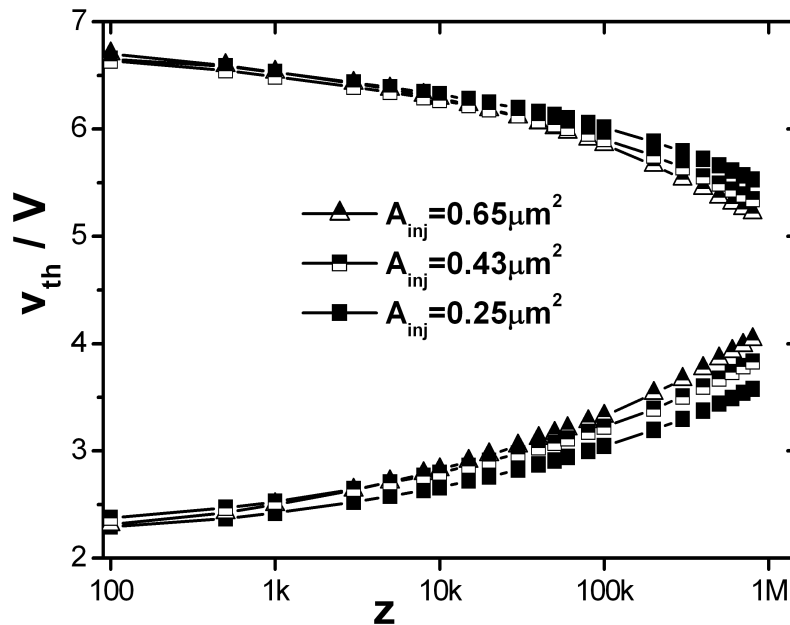


Bild 3.20: Darstellung der Schwellspannung als Funktion der Zykelzahl für Speicherzellen mit unterschiedlich großen Injektorflächen.

4 Untersuchungen zur Oxiddegradation

4.1 Bauelemente der Single-Poly-Speicherzellen

Die Single-Poly-Speicherzelle besteht aus mehreren Teilkomponenten. Hierzu zählen der Select-Transistor, der zur Selektion der Speicherzelle innerhalb einer Arraystruktur verwendet wird und der Auslesetransistor, siehe Bild 4.1. Weitere Komponenten sind die Control-Gate-Kapazität und der EEPROM-Injektor, von dessen Qualität der Dateinerhalt und die Datenwechselstabilität maßgeblich abhängen.

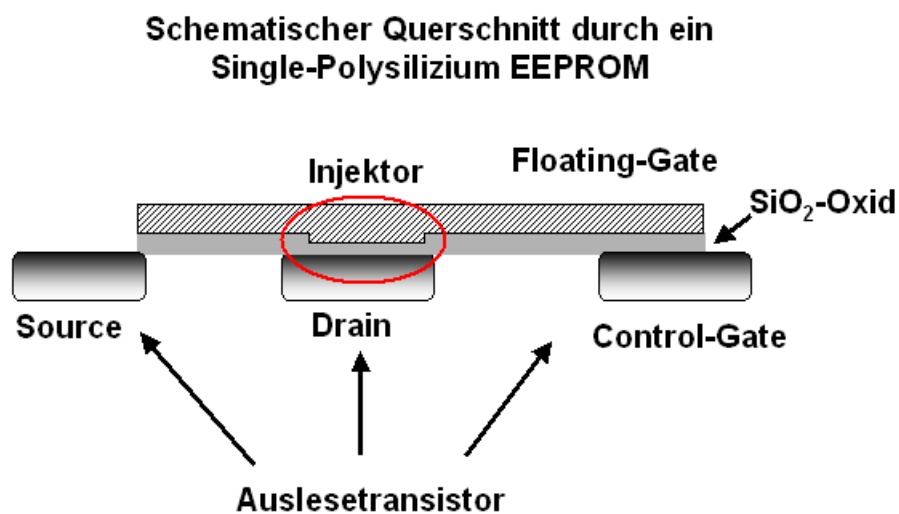


Bild 4.1: Zu den wichtigsten Bauelementen der Single-Polysilizium EEPROM zählen der Auslesetransistor, der Select-Transistor, die Control-Gate-Kapazität und der Injektor.

4 Untersuchungen zur Oxiddegradation

4.1.1 Auslesetransistor

Der Auslesetransistor der Single-Poly-Speicherzellen besteht aus Source, Drain und einem Gate (Polyfinger 2), das über die n-Wanne des Control-Gates steuerbar ist. Im Gegensatz zu Flash-Speicherzellen findet die Programmierung und Löschung der Zelle nicht über den Kanal des Auslesetransistors statt, sondern erfolgt hiervon getrennt über den Injektor. Die beim Programmieren und Löschen erzeugten Defektzentren und Oxidladungen haben daher keinen Einfluss auf die Schwellspannungen des Transistors.

Für die Bestimmung von Schwellspannungen wird, wie in Kapitel 3 beschrieben, die Control-Gate-Spannung so lange variiert, bis ein Drainstrom von $5 \mu\text{A}$ erreicht wird. Die so ermittelte Control-Gate-Spannung wird als Schwellspannungswert ausgegeben. Da sich zwischen Substrat und n-Wanne ein pn-Übergang befindet, kann das Control-Gate nicht beliebig negativ polarisiert werden, da sich bei einer bestimmten Spannung die Diode zwischen n-Wanne und Substrat öffnet. Dies geht deutlich aus den Transferkennlinien des Auslesetransistors hervor, Bild 4.2 a). Ab einer Control-Gate-Spannung

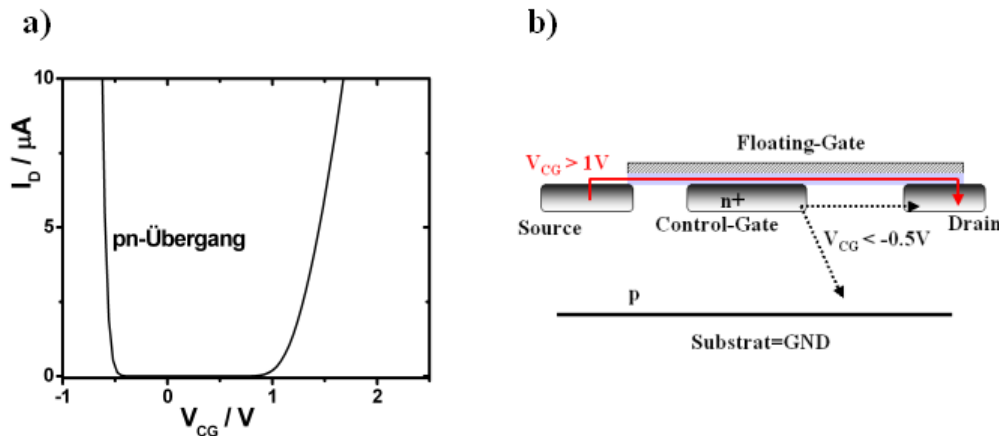


Bild 4.2: a) Transferkennlinie des Auslesetransistors: Der Drainstrom I_D wurde über der Control-Gate Spannung aufgetragen ($V_D=2 \text{ V}$ und $V_S=0 \text{ V}$). b) Schematische Darstellung des Zellquerschnitts. Dargestellt ist der Stromfluß für $V_{CG} > 1 \text{ V}$ und $V_{CG} < -0.5 \text{ V}$

von 1 V wird der Auslesetransistor leitend, wodurch Schwellspannungen im programmierten Zustand messbar sind. Bei Gatespannungen kleiner -0.5 V öffnet sich die Diode zwischen Control-Gate und Substrat, wodurch der Drainstrom ansteigt und Schwellspannungen im gelöschten Zustand $\leq -0.5 \text{ V}$ nicht messbar sind. Durch die Öffnung des pn-Überganges zwischen Control-Gate und Substrat können die in das Substrat injizierten Ladungsträger über den Substrat- und Drainkontakt abfließen, siehe Bild 4.2b).

Aus diesem Grund wird bei Single-Poly-Speicherzellen mit einer negativen Substratspannung gemessen, um auch „tief“ liegende Schwellspannungen des gelöschten Zustandes messen zu können. Die negative Substratspannung führt aufgrund des Substratsteuereffektes [24] zu einer Verschiebung der Schwellspannungen bzw. der Transferkennlinien. Um den Substratsteuereffekt messen zu können, wurden an einem einzelnen

Auslesetransistor Transferkennlinien bei einer von null verschiedenen Substrat-Source aufgenommen. In Bild 4.3 sind diese für verschiedene Substrat-Source-Spannung $V_{BS} = -1, -2, -3, -4$ und -5 V dargestellt. Die Transferkennlinien mit Substratspannung sind deutlich gegenüber der spannungsfreien Kennlinie verschoben. Die Kennlinienschar zeigt, dass die Transferkennlinie bei einer Substratspannung von $V_{BS} = -5$ V um 1.6 V gegenüber der spannungsfreien Messung verschoben ist. Das bedeutet, dass bei einer Substratspannung von -5 V die Schwellspannungen im programmierten Zustand um 1.65 V zu verringern und diejenigen im gelöschten Zustand um 1.65 V anzuheben sind.

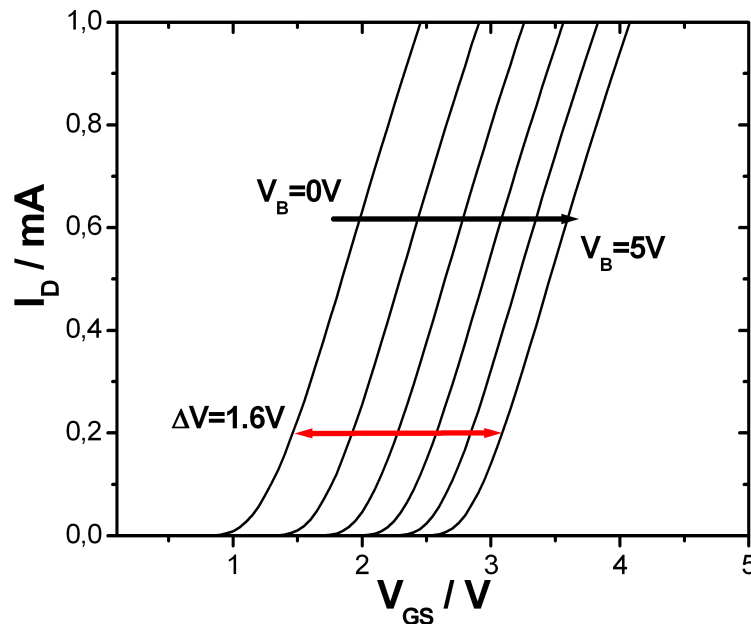


Bild 4.3: Zeigt die Transferkennlinien des Auslesetransistors bei verschiedenen Substratspannungen. Der Drainstrom wurde in Abhängigkeit von der Control-Gate Spannung gemessen.

4.1.2 Select-Transistor

Der Select-Transistor hat die Funktion, die Speicherzelle innerhalb eines Arrays zu selektieren, um z. B. Programmier- und Löschpulse richtig adressieren zu können, siehe Bild 4.1. Beim Löschen der Speicherzelle wird die Löschspannung über den n-Kanal des Select-Transistors unter den EEPROM-Injektor geleitet. Bei diesem Vorgang findet ein Spannungsabfall über dem n-Kanal des Select-Transistors statt, der die Löschspannung am Injektor verringert. Im Löschfall liegen am Drainkontakt und Gatekontakt des Select-Transistors die volle Löschspannung V_P an. Damit über den Transistor ein Löschstrom von ca. 1 nA fließen kann, muss ein leitfähiger n-Kanal ausgebildet sein. Daher muss die Spannungsdifferenz zwischen Source (Injektorgebiet) und Gate größer gleich der Schwellspannung v_{th} des Transistors sein. Da der Sourcekontakt des Select-Transistors auf positivem Potenzial liegt und somit eine negative Spannung $V_{SB} < 0$ zwischen Substrat und Sourcekontakt anliegt, wird die Schwellspannung des Select-Transistors aufgrund des Substratsteuereffektes angehoben. Für den Spannungsabfall

4 Untersuchungen zur Oxiddegradation

ergibt sich folgende Relation: $\Delta V_{SG} \geq v_{th}(V_{SB}) \geq v_{th}(0)$. Um diesen Spannungsabfall zu verifizieren, wurde an einem MOS-Transistor die Löschspannung V_P an Gate und Drain angelegt und ein definierter Strom in den Source-Anschluss eingeprägt, siehe Bild 4.4. Bei einem Strom von ca. 1 nA wird der Transistor gerade leitend. Die ge-

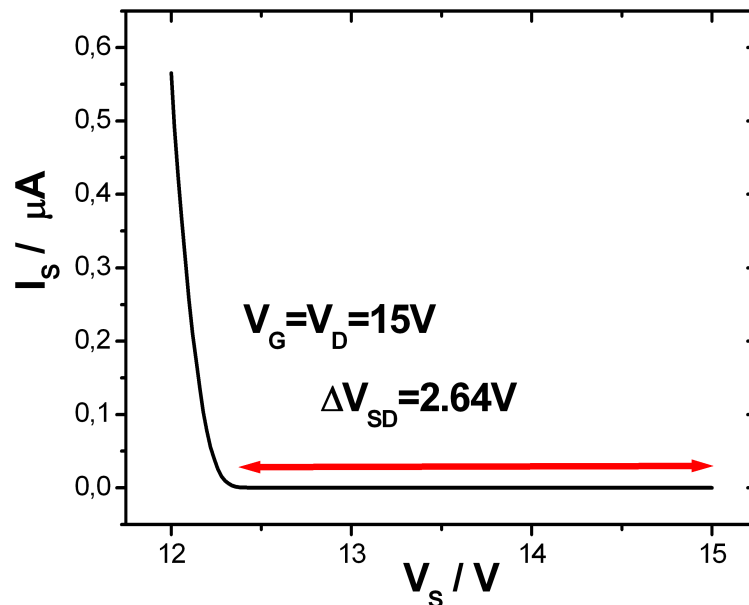


Bild 4.4: Zeigt den Source-Strom I_S eines MOS-Transistors in Abhängigkeit von der Sourcespannung V_S . Die Drain- und Gatespannung lag bei 15 V. Bei einem Strom von 1 nA wurde eine Source/Drain-Spannung von 2.64 V gemessen.

messene Source-Drain-Spannung beträgt $\Delta V_{SD} = 2.64$ V und liegt deutlich über der Schwellspannung des hier verwendeten (natural) Transistors, die bei $v_{th}(0)=0$ V liegt.

Messungen am Select-Transistor der Single-Poly-Speicherzellen liefern einen Spannungsabfall von 1.7 V [86], sodass am EEPROM-Injektor bei $V_P=15$ V eine effektive Löschspannung von 13.3 V anliegt.

4.1.3 Injektor

Das Layout der hier betrachteten EEPROM-Injektoren ist schematisch in Bild 4.5a) zu sehen. Der Injektor befindet sich im Aktivgebiet und besteht im Wesentlichen aus einem Polysilizium-Streifen, einer Tunneloxidschicht zwischen Polysilizium und Substrat sowie einer Injektorimplantation. Der Querschnitt entlang der Schnittlinie in Bild 4.5a) ist in Bild 4.5b) zu sehen und zeigt das Tunneloxid, das Polysilizium und die n-Wanne der Injektorimplantation. Aufgrund der geringen Größe eines einzelnen Injektors liegen die zu messenden Ströme im pA-Bereich, weshalb eine Teststruktur verwendet wird, die aus 20 x 40 parallel geschalteten Injektorstrukturen besteht und Ströme im nA-Bereich liefert, Bild 4.6. Hierüber kann für die gate- und substratseitige Injektion der mittlere Strom pro Injektor bestimmt werden. Die Injektoren der Teststruktur sind baugleich mit den EEPROM-Injektoren und wurden mit demselben Prozess wie die Speicherzelle

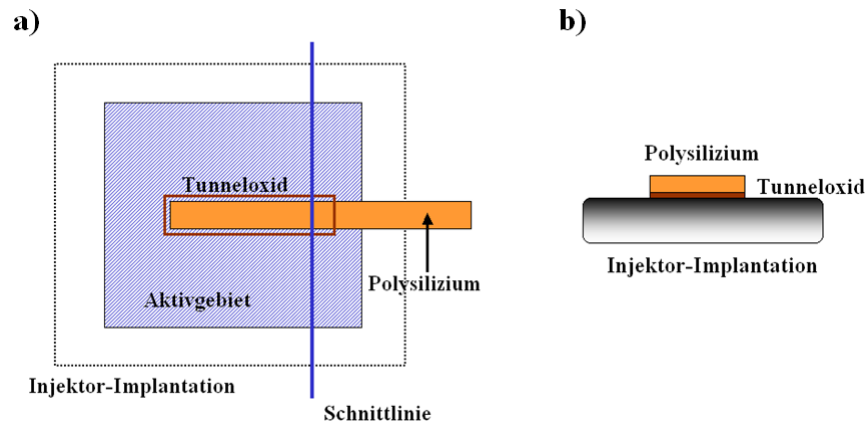


Bild 4.5: a) Schematisches Layout eines einzelnen Injektors: Die wichtigsten Injektorelemente sind das Aktivgebiet, das Polysilizium, das Tunneloxid sowie die Injektorimplantation. b) Zeigt den Querschnitt durch das Layout der Injektorstruktur entlang der Schnittlinien in Bild 4.5a).

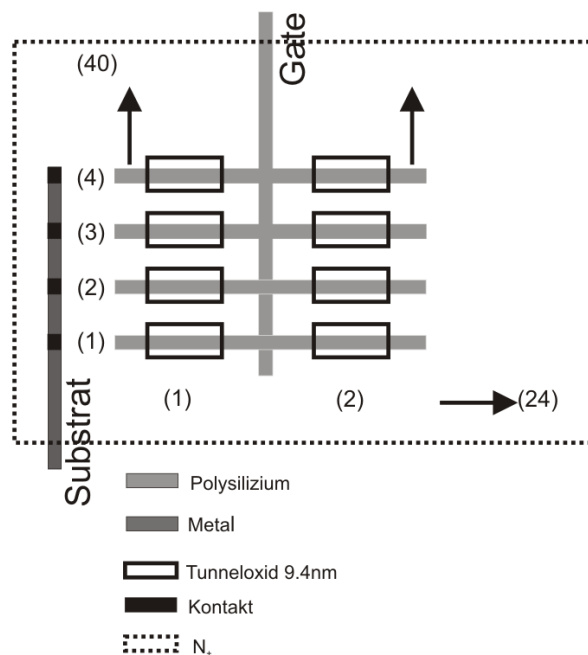


Bild 4.6: Schematische Darstellung der Teststruktur. Die Teststruktur setzt sich aus 24x40 parallel geschalteten Injektoren zusammen, die baugleich mit den EEPROM Injektoren sind.

4 Untersuchungen zur Oxiddegradation

entwickelt. Eine solche Teststruktur bietet den Vorteil, dass die FN-Injektion unter den gleichen Bedingungen stattfindet wie in der Speicherzelle. Würde eine einzelne, großflächige Kapazität verwendet, so kann nicht ohne Weiteres auf die FN-Injektion im EEPROM-Injektor geschlossen werden, da die Fowler-Nordheim-Injektion geometrieabhängig sein kann und z. B. von der Kantenanzahl beeinflusst wird [92]. Messungen

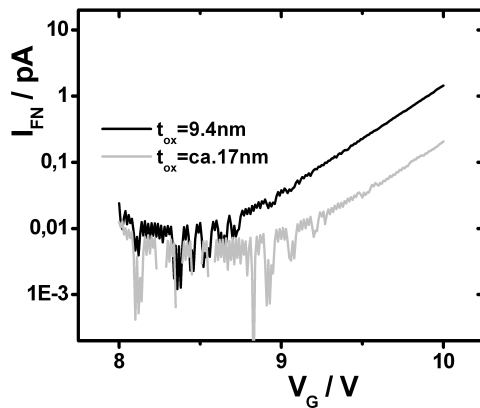


Bild 4.7: Der Fowler-Nordheim-Strom wurde in Abhängigkeit von der Gatespannung gemessen (Fowler-Nordheim Kennlinie).

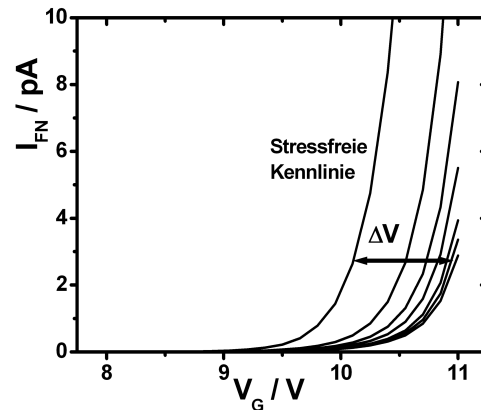


Bild 4.8: Verschiedene Fowler-Nordheim Kennlinien nach elektrischem Stress.

zu den mittleren Injektorströmen finden sich in Bild 4.7. Der Fowler-Nordheim-Strom wurde an Testkapazitäten mit unterschiedlichen Oxiddicken gemessen und über der Gatespannung aufgetragen. Der Strom als Funktion der Gatespannung weist auf eine deutliche Abnahme des FN-Stromes mit zunehmender Oxiddicke hin. Um unterschiedliche Oxiddicken zu realisieren, wurde das Siliziumdioxid über verschieden hoch dotiertem Substrat aufgewachsen [93,94]. Bei der verwendeten n_- -Implantation von etwa 10^{17} cm^{-3} betrug die Oxiddicke 9.4 nm, wohingegen über der n_+ Implantation von etwa 10^{21} cm^{-3} eine Oxiddicke von ca. 17 nm erreicht wurde. Die FN-Injektion wird in diesem Fall von der Oxiddicke und der Dotierstoffkonzentration unterhalb des Injektor bestimmt. Die höhere n_- -Implantation führt zu einer höheren Oxiddicke, was eine unmittelbare Reduktion des elektrischen Feldes bzw. des FN-Stromes bewirkt. Die Dotierstoffkonzentration hat einen hierzu gegenläufigen Effekt zur Folge, da mit steigender Dotierstoffkonzentration das Fermi-niveau ansteigt und sich hierüber die Barrierenhöhe ϕ verringert. Die hierdurch bedingte Erhöhung der Fowler-Nordheim-Injektion fällt, wie aus Bild 4.7 hervorgeht, deutlich geringer aus als die durch die Zunahme der Oxiddicke verursachte Abnahme der FN-Injektion.

Der Einfluss von Oxidladungen auf den Fowler-Nordheim-Strom geht aus Bild 4.8 hervor. Hierzu wurde eine Injektorstruktur schrittweise durch elektrischen Stress gealtert und der FN-Strom in Abhängigkeit von der Gatespannung gemessen. Der elektrische Stress führt zu einer deutlichen Spannungsverschiebung ΔV der FN-Kennlinien. Die durch den elektrischen Stress erzeugten Oxidladungen verringern das elektrische Feld über dem Tunneloxid deutlich, was die FN-Kennlinien zu höheren Spannungen hin verschiebt. Die Oxidladungen haben zusätzlich die Erhöhung der Potenzialbarriere zwi-

schen Silizium und Siliziumdioxid Φ zur Folge, was sich unmittelbar auf die Tunnelkonstanten auswirkt [95,96]. Die Zunahme bzw. Abnahme der A - und B -Konstante mit steigender Barrierenhöhe kann anhand der Formeln für die Tunnelkonstanten in Kapitel 2 verstanden werden. Für das Änderungsverhalten der Tunnelkonstanten gelten folgende Relationen: $A = \frac{e^3 m_e}{8\pi m_{ox} \Phi} \Rightarrow A \downarrow \text{ wenn } \phi \uparrow$ $B = \frac{4\sqrt{2m_{ox}\Phi^3}}{3he} \Rightarrow B \uparrow \text{ wenn } \phi \uparrow$

Die Spannungsverschiebung ΔV hängt linear von der eingebauten Oxidladung ab, wodurch sich hierüber die Oxidladung bestimmen lässt. Das hierauf basierende Messverfahren soll im nächsten Kapitel detailliert vorgestellt werden.

4.2 DiMaria-Methode zur Bestimmung der Oxidladung

Der Einbau von Oxidladungen ist die Ursache für die Abnahme des Schwellspannungsfensters. Aus diesem Grund sollen Messverfahren eingeführt werden, um die Oxidladung in Abhängigkeit von Stressparametern untersuchen zu können. Der Zusammenhang zwischen Oxidladung und injizierter Ladung ist bereits in Kapitel 2 theoretisch untersucht worden und soll auch im Vordergrund der experimentellen Untersuchungen stehen. Der Ladungseinbau im Tunneloxid der Injektorstrukturen wird mit der DiMaria-Methode [97-99] untersucht. Diese Methode ermöglicht die Messung von Oxidladungen und Ladungsschwerpunkten in SiO_2 und basiert auf der Fowler-Nordheim-Spannungsverschiebung. Ein im I-V Diagramm definierter Arbeitspunkt (V_A, I_A) wird durch den Einfluss der Oxidladungen zu höheren Spannungen hin verschoben, Bild 4.9. Die Span-

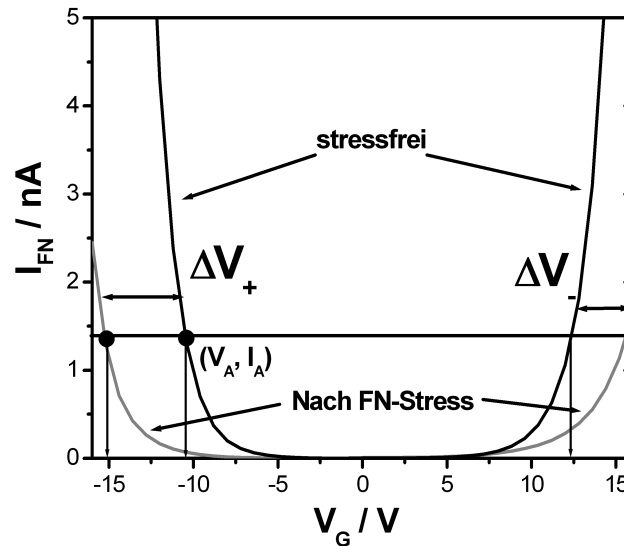


Bild 4.9: Verschiebung der FN-Kennlinien aufgrund von Oxidladungen. Bei einem Auslestestrom I_g wird die korrespondierende Spannung gemessen und die Spannungsverschiebungen relativ zur „stressfreien“ Kennlinien berechnet.

nungsverschiebung ist abhängig von der Injektionsrichtung und unabhängig vom Strom I_A . Die Größe ΔV_+ bezeichnet im Weiteren die Spannungsverschiebung nach substrat-

4 Untersuchungen zur Oxiddegradation

seitiger Injektion und ΔV_- die nach gateseitiger Injektion. Die Größe ΔQ_{ox} bezeichnet die durch den elektrischen Stress hervorgerufenen Änderungen der Oxidladung. Für die Bestimmung von ΔQ_{ox} wird vorausgesetzt, dass sich die Oxidladung durch eine δ -Verteilung mit dem Ladungsschwerpunkt x beschreibbar ist. Die Position des Ladungsschwerpunktes wird relativ zur Substratoberfläche gemessen.

Die Spannungen über dem Tunneloxid V_{ox} müssen unter Berücksichtigung der Oxidladungen neu berechnet werden, was mithilfe der Poissonschen Gleichung und dem Gaußschen Satz geschieht. Eine einfache Rechnung liefert Gleichungen für die Spannungen bei substratseitiger und gateseitiger Injektion:

$$a) \quad V_{ox} = V_G - \frac{\Delta Q_{ox}}{\epsilon_{Si}\epsilon_0}(t_{ox} - x) \quad \Rightarrow \quad b) \quad \Delta V_- = \frac{\Delta Q_{ox}}{\epsilon_{Si}\epsilon_0}(t_{ox} - x) \quad (4.1)$$

$$a) \quad V_{ox} = V_S - \frac{\Delta Q_{ox}}{\epsilon}x \quad \Rightarrow \quad b) \quad \Delta V_+ = \frac{\Delta Q_{ox}}{\epsilon}x \quad (4.2)$$

Aus den berechneten Spannungsverschiebungen Gl.4.1b) und Gl.4.2b) kann die Änderung der Oxidladung und der relative Ladungsschwerpunkt x_R ($x_R = x/t_{ox}$) berechnet werden, was folgende Gleichungen liefert.

$$a) \quad \Delta Q_{ox} = \frac{\epsilon}{t_{ox}}(\Delta V_+ + \Delta V_-) \quad b) \quad x_R = \frac{\Delta V_+}{\Delta V_+ + \Delta V_-} \quad (4.3)$$

Wie in [41,100] gezeigt wurde, sind diese Gleichungen exakt, wenn der Ladungsschwerpunkt außerhalb der Tunnelstrecke lokalisiert ist. Befindet er sich innerhalb der Tunnelstrecke (z. B. bei dünnen Oxiden), ist der Tunnelstrom nicht mehr durch das Fowler-Nordheim-Gesetz beschreibbar, sodass der lineare Zusammenhang zwischen Spannungsverschiebung und Ladungsschwerpunkt verloren geht. In diesem Fall muss die DiMaria-Methode durch geeignete Korrekturgrößen modifiziert werden [100]. Alternativ kann auch ein anderes Verfahren eingesetzt werden, das nicht die Spannungsverschiebung, sondern die Fowler-Nordheim-Ableitung betrachtet [101,102].

Für die Bestimmung weiterer Oxidgrößen, wie z. B. der Oberflächenladung Q_{ss} und Grenzflächendefektdichte N_{ss} , kann die DiMaria-Methode mit anderen Messverfahren kombiniert werden. So findet sich bei Y. Nissan-Cohen et al. [98] eine Kombination aus DiMaria-Methode und C-V-Messungen, um Oberflächendefekte bestimmen zu können. Ferner kann hierfür auch die Charge-Pumping (CP) Technik [103,100] eingesetzt werden. Für den Fall, dass die Messungen an einem MOSFET erfolgen, kann über die Transferkennlinien („Subthreshold“-Bereich) des MOSFET die Oberflächendefektdichte und Oberflächenladung gemessen werden [104,105].

Die hier verwendeten Oxide mit einer Dicke zwischen 9 und 10 nm können mit der DiMaria-Methode charakterisiert werden, die die gesamte im Oxid eingefangene Oxidladung mißt. Eine weitere Differenzierung in Oberflächen- und Volumenoxidladung ist nicht notwendig.

4.2.1 Stressmessungen

Beim Spannungsstress wird zwischen konstantem Spannungsstress „Constant Voltage Stress“ (CVS) und gepulstem Spannungsstress, „Pulsed Voltage Stress“ (PVS) unterschieden. Ferner gibt es noch den konstanten Stromstress „Constant Current Stress“ (CCS). Weitere Stressformen sind die Strom- und Spannungsrampen „Ramp Voltage Stress“ oder Stromrampen mit exponentiellem Verhalten („Exponentially Ramped Current Stress“) [106,107].

Beim CVS wird am Gate- oder Substratkontakt des Bauteiles eine konstante Spannung angelegt und der korrespondierende Strom gemessen. Unter CCS wird dem Oxid eine konstante Stromdichte eingeprägt und der Spannungsabfall gemessen. In beiden Fällen (CVS und CCS) wird die injizierte Ladung durch Integration der Ströme ermittelt. Für die Anwendung der FN-Methode musste der CCS und CVS in vordefinierten Zeitintervallen unterbrochen werden, um die Fowler-Nordheim-Spannungsverschiebung messen zu können. Beim konstanten Spannungsstress wurden bei einem Strom von $I_A = 2 \text{ nA}$ die jeweiligen Spannungswerte und Spannungsverschiebungen bestimmt. In Bild 4.10 sind die Spannungsverschiebungen (linke Achse) und der relative Ladungsschwerpunkt in Abhängigkeit von der Stresszeit dargestellt. Die Oxidladung und die injizierte Ladung sind in Bild 4.11 zu sehen. Das sublineare Verhalten der injizierten Ladung ist deutlich erkennbar. Ähnliche Ergebnisse wurden mit bipolarem CVS erzielt.

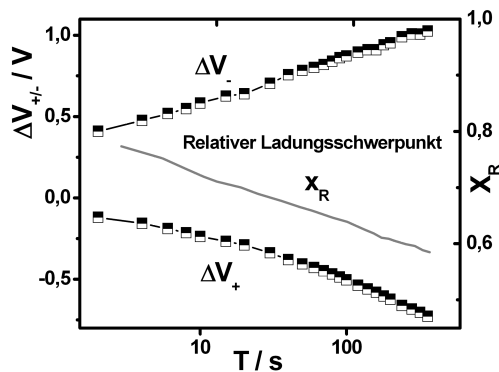


Bild 4.10: Der relative Ladungsschwerpunkt (rechte Achse) sowie die Spannungsverschiebungen wurden in Abhängigkeit von der Stresszeit gemessen ($V_G = 11.75 \text{ V}$).

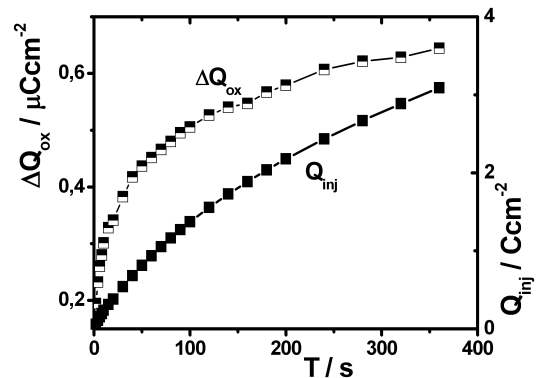


Bild 4.11: Die Oxidladung (linke Achse) und die injizierte Ladung (rechte Achse) sind über der Stresszeit aufgetragen.

Am Gate- und Substratkontakt lag abwechselnd eine Spannung von 12 V an. Dadurch konnte die Degradation in beide Injektionsrichtungen erfolgen. In Bild 4.12 ist hierzu die Oxidladung (linke Achse) und die injizierte Ladung (rechte Achse) als Funktion der Stresszeit T dargestellt. Die injizierte Ladung setzt sich aus einem substrat- und einem gateseitigen Beitrag zusammen. Die substratseitige Injektion fällt höher aus als die gateseitige Injektion, was vermutlich auf die unterschiedlich hohe Rauigkeit der Grenzfläche zurückzuführen ist. Die Grenzfläche zwischen Polysilizium und Oxid weist eine höhere Rauigkeit auf als die Grenzfläche zwischen Substrat und Oxid. Aufgrund der Unebenheiten sind über einer rauen Grenzfläche die (mittleren) elektrischen Fel-

4 Untersuchungen zur Oxiddegradation

der größer als bei einer vergleichsweise glatten Grenzfläche. Die Folge hiervon ist eine richtungsabhängige Fowler-Nordheim-Injektion [6,108,109]. Eine weitere Ursache könnte die asymmetrische Lage des Ladungsschwerpunktes sein, wodurch die Stärke des elektrischen Feldes und damit die Fowler-Nordheim-Injektion richtungsabhängig wären. In Bild 4.13 wurde die Oxidladung unter CCS bestimmt und eine Stromdichte von $j_0 = \pm 100 \text{ mA cm}^{-2}$ verwendet. Für die Bestimmung der Spannungsverschiebung wurde der CCS unterbrochen, dem Bauteil ein Strom von $\pm 13 \text{ mA}$ eingeprägt und die korrespondierende Spannung gemessen. Aus den berechneten Spannungsverschiebungen ließ sich die Oxidladung für beide Injektionsrichtungen berechnen. Die injizierte Ladung (rechte Achse) folgt aus der Relation $Q_{inj} = j_s \cdot T$ und zeigt ein einfaches lineares Verhalten.

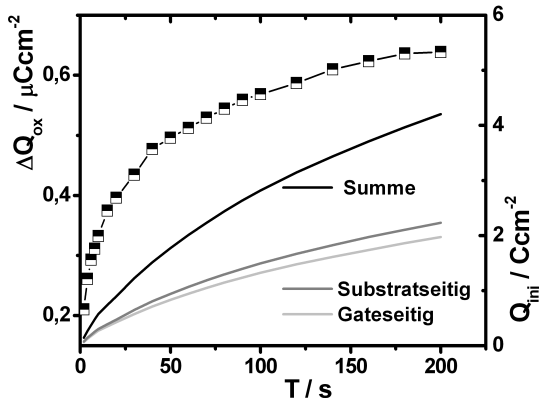


Bild 4.12: Die Testkapazität wurde bipolar ($\pm 12 \text{ V}$) gestresst. Die Oxidladung (linke Achse) und die injizierte Ladung wurden über der Stresszeit T aufgetragen.

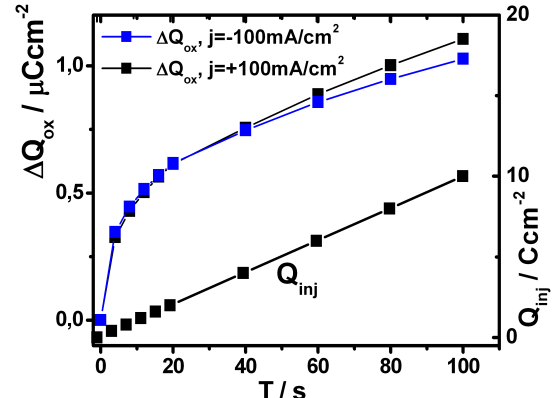


Bild 4.13: Ladungseinbau unter CCS: Die Oxidladung (linke Achse) wurde für beide Injektionsrichtungen gemessen und als Funktion der Stresszeit dargestellt.

4.2.2 Einfluss von Oxiddicke und Temperatur

Es soll nachfolgend die Oxidladung als Funktion der injizierten Ladung gemessen werden, um die Temperatur- und Oxiddickenabhängigkeit des Ladungseinbaus zu verifizieren. Der funktionale Zusammenhang zwischen Oxidladung und injizierter Ladung $\Delta Q_{ox}(Q_{inj})$ soll im Mittelpunkt der Untersuchungen stehen. Es wurden verschiedene Oxiddicken untersucht und die Oxidladung und der relative Ladungsschwerpunkt in Abhängigkeit von der injizierten Ladung gemessen. In Bild 4.14 ist die Oxidladung für drei unterschiedliche Oxiddicken dargestellt. Der relative Ladungsschwerpunkt x_R ist in Bild 4.15 zu sehen und verschiebt sich mit zunehmender Oxiddicke in Richtung Oxidmitte. Es zeigt sich, dass der Trapping-Yield mit abnehmender Oxiddicke sinkt und die Oxidalterung verlangsamt wird. In diesem Zusammenhang kann für den Einfang von negativen Oxidladungen eine untere Grenzdicke x_{DF} des Oxides angegeben werden [110]. Bei Tunneloxiden mit $t_{ox} \leq x_{DF}$ ist der Trapping-Yield so gering, dass

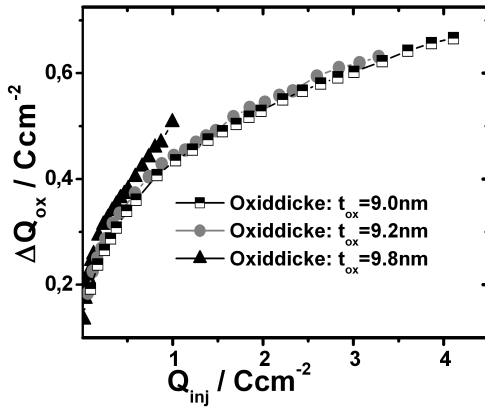


Bild 4.14: Für verschiedene Oxiddicken wurde die Oxidladung in Abhängigkeit von der injizierten Ladung gemessen.

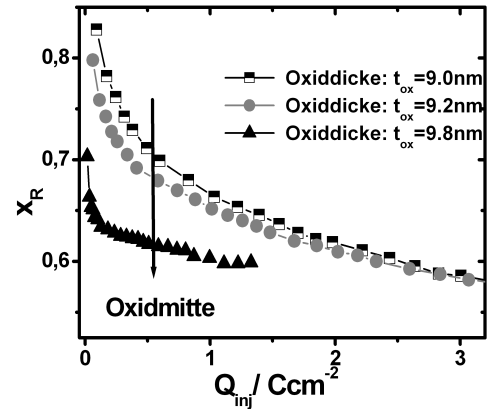


Bild 4.15: Der relative Ladungsschwerpunkt x_R wird mit steigender Oxiddicke in Richtung Oxidmitte verschoben.

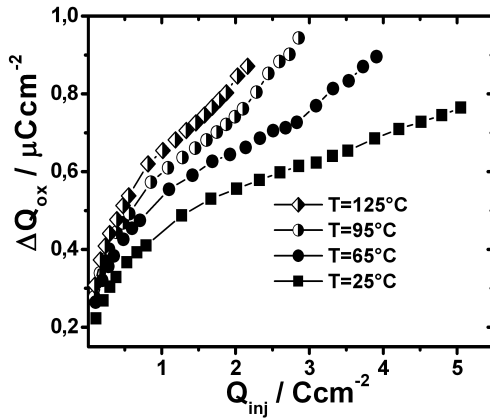


Bild 4.16: Für verschiedene Temperaturen wurde die Oxidladung als Funktion der injizierten Ladung dargestellt.

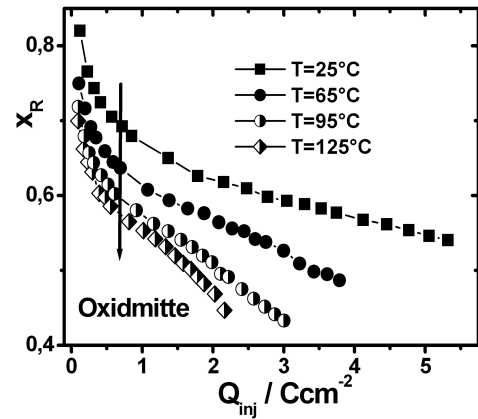


Bild 4.17: Darstellung des relativen Ladungsschwerpunktes $x(Q_{inj})$ für verschiedene Temperaturen.

hiervon abhängige Größen, wie z. B. der K-Parameter des Papadas-Modelles, vernachlässigbar klein sind $K \approx 0$. Für die Abschätzung der Grenzdicke wird angenommen, dass die Ladungsträger während des Tunnelvorganges keine Defekte erzeugen und außerhalb der Tunnelstrecke erst nach dem Zurücklegen der mittleren freien Weglänge x_{fW} Defekte und Oxidladungen generieren. Der quasi defektfreie Bereich x_{DF} kann näherungsweise durch die Summe aus Tunnelstrecke x_T und mittlerer freier Weglängen x_{fW} beschrieben werden: $x_{DF} \approx x_T + x_{fW}$. In der Literatur finden sich für die mittlere freie Weglänge von Ladungsträgern in SiO_2 Werte zwischen 1.5 und 3 nm [111,112]. Diese Werte liefern mit einer FN-Tunnelstrecke von etwa 3nm eine Grenzdicke zwischen 4.5 und 6 nm, bei der der Einbau von negativen Oxidladungen vernachlässigbar ist. Eine weitere Untersuchung betraf die Temperaturabhängigkeit des Ladungseinbaues. Für verschiedene Temperaturen wurde ΔQ_{ox} und x_R gemessen und in Abhängigkeit von der injizierten Ladung dargestellt. In Bild 4.16 ist eine deutliche Zunahme

4 Untersuchungen zur Oxiddegradation

der Oxidladung mit steigender Temperatur zu beobachten, was auf einen Anstieg der Defekterzeugung und Defektbesetzung hindeutet. Der Ladungsschwerpunkt wird mit zunehmender Temperatur in Richtung Oxidmitte verschoben, Bild 4.17. Ein möglicher Mechanismus, der diese Beobachtungen erklären könnte, basiert auf der Temperaturabhängigkeit der Elektronenenergie. Mit steigender Temperatur nimmt die Anzahl der Elektronen mit hoher kinetischer Energie zu, was aus der temperaturbedingten Verbreiterung der Energieverteilung (Fermiverteilung) hervorgeht. Mit zunehmender Elektronenenergie können Bindungen leichter aufgebrochen werden, sodass z. B. Dangling-Bond-Defekte entstehen. Das Aufbrechen von Bindungen könnte auch die Freisetzung von Wasserspezies bewirken, was eine zusätzliche Oxidschädigung nach sich ziehen würde [134]. Ein weiterer möglicher Effekt für die höhere Oxiddegradation könnte die Verkürzung der Tunnelstrecke mit steigender Elektronenenergie sein. Wie in Bild 4.18 schematisch gezeigt wird, verkürzt sich die Tunnelstrecke mit steigender Elektronenenergie bzw. Temperatur. Damit verlängert sich die Wegstrecke durch das Oxid, auf der Oxidschädigungen stattfinden können, sodass die Wahrscheinlichkeit für die Erzeugung und Besetzung von Defekten zunimmt [113].

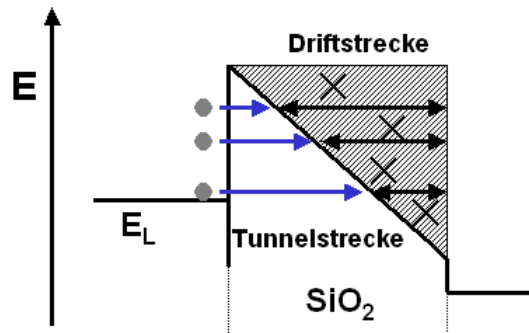


Bild 4.18: Bandschema für das Fowler-Nordheim-Tunneln: Für Elektronen mit hoher kinetischer Energie nimmt die Tunnelstrecke x_T ab. Durch die Verkürzung der Tunnelstrecke nimmt die Driftstrecke, auf der Oxidschädigungen stattfinden können, zu.

4.3 Ladungseinbau in Speicherzellen (Papadas-Methode)

Die Oxidladung im Tunneloxid einer Speicherzelle lässt sich mit der DiMaria-Methode nicht extrahieren und muss über ein anderes Verfahren bestimmt werden, das auf der theoretischen Beschreibung der Datenwechselstabilität basiert. Für die Anwendung der DiMaria-Methode wäre es notwendig, das Floating-Gate zu kontaktieren, um die Fowler-Nordheim-Spannungsverschiebung zu messen. Eine so modifizierte Speicherzelle hätte ein völlig anderes elektrisches Verhalten als die ursprüngliche Speicherzelle, da der zusätzliche Floating-Gate-Kontakt und die damit verbundenen Anschlüsse die Kapazität zwischen Floating-Gate und Substrat verändern würden. Es muss daher versucht werden, die Oxidladung direkt aus der Charakteristik der Datenwechselstabilität zu bestimmen, was einen analytischen Ausdruck zwischen Schwellspannungen, Layoutparametern und der Oxidladung erfordert. Dieser soll im nächsten Schritt abgeleitet

Ersatzschaltbild des Injektors

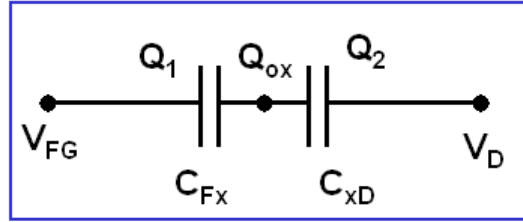


Bild 4.19: Ersatzschaltbild für den EEPROM-Injektor mit Oxidladung. Die Injektor-kapazität C_{inj} setzt sich aus den Ersatzkapazitäten C_{Fx} und C_{xD} zusammen.

werden und findet sich in [25,114]. Für die Herleitung wird eine Speicherzelle betrachtet, die analog zur Single-Poly-Speicherzelle programmiert und gelöscht wird. In Bild 4.19 ist das Ersatzschaltbild des EEPROM-Injektors mit eingebauter Oxidladung zu sehen. Wie bereits in Kapitel 3 gezeigt wurde, kann mithilfe des Ladungserhaltungssatzes ein Zusammenhang zwischen den Oxidladungen und den Spannungsgrößen hergeleitet werden. Für die Oxidladung Q_1 lässt sich folgender Ausdruck ableiten,

$$Q_1 = C_{inj}(V_{FG} - V_D) - Q_{ox} \frac{C_{inj}}{C_{xD}} \quad \text{mit} \quad C_{inj} = \frac{C_{Fx}C_{xD}}{C_{Fx} + C_{xD}}, \quad (4.4)$$

wobei die Injektor-kapazität C_{inj} als Gesamtkapazität des Spannungsteilers definiert ist, V_D die Drainspannung und V_{FG} das Floating-Gate-Potenzial bezeichnet.

$$a) \quad V_{FG}(t) \propto Q_{FG}(t) \Rightarrow b) \quad \frac{dV_{FG}}{dt} = \frac{1}{C_T} \frac{dQ_{FG}}{dt} \quad (4.5)$$

Aus der Relation zwischen Floating-Gate-Potenzial und Floating-Gate-Ladung Gl. 4.5 a) folgt durch Differenzieren Gl. 4.5 b). Hierbei steht C_T für die Gesamtkapazität der Speicherzelle. Wenn vorausgesetzt wird, dass der Programmier- und Löschpuls unendlich steile Flanken besitzen (Rechteckpulse), kann nach einfachen Umformungen ein Zusammenhang zwischen Floating-Gate-Ladung und der maximalen und minimalen elektrischen Feldstärke E_{min} und E_{max} hergestellt werden, Gl. 4.6. Das minimale und maximale elektrische Feld bezieht sich auf die Spannung über dem Tunneloxid, siehe Bild 4.20. Es gilt:

$$Q_{FG}^{P/E} = \frac{A_{inj}\epsilon_0\epsilon_{Si}}{k_D}(E_{max} - E_{min}), \quad (4.6)$$

wobei k_D der in Kapitel 2 eingeführte Drain-Kopplungsfaktor ist, der sich über die Injektor-kapazität und die Gesamtkapazität berechnen lässt $C_{inj} : C_T = k_D$.

Wie die Diskussion in Kapitel 3.3.2 gezeigt hat, beeinflussen die Oxidladungen lediglich die maximale elektrische Feldstärke E_{max} , wohingegen die minimale Feldstärke E_{min} unabhängig von der eingebauten Ladung ist [115]. Damit vereinfacht sich 4.6 zu

4 Untersuchungen zur Oxiddegradation

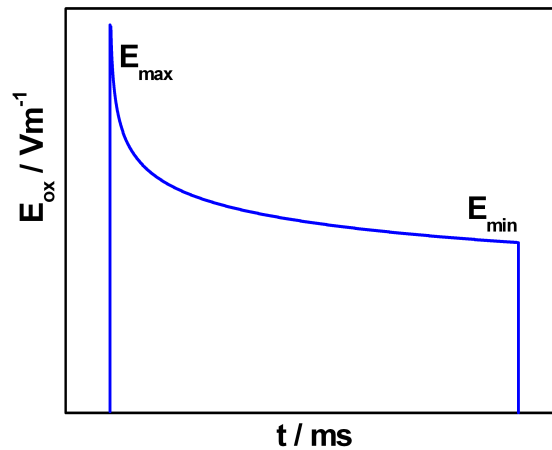


Bild 4.20: Darstellung des elektrischen Feldes für einen Spannungspuls mit unendlich steilen Anstiegs- und Abfallsflanken. E_{max} und E_{min} bezeichnen das Maximum bzw. Minimum.

4.7 a), woraus unmittelbar Gl. 4.7 b) folgt

$$a) \quad Q_{FG}^{P/E} = \frac{A_{inj} \epsilon_0 \epsilon_{Si}}{k_D} E_{max} \Rightarrow b) \quad \Delta Q_{FG}^{P/E} = \frac{A_{inj} \epsilon_0 \epsilon_{Si}}{k_D} \Delta E_{max} \quad (4.7)$$

Die Änderung der maximalen Feldstärke kann über den gaußschen Satz mit der Änderung der Oxidladung verbunden werden, was Gleichung 4.8 liefert.

$$\Delta E_{max}^{E/P} = -\frac{1}{A_{inj} \epsilon} \frac{dQ_{1,2}}{dQ_{ox}} \Delta Q_{ox} \quad (4.8)$$

Wird die Gleichung 4.4 mit 4.8 kombiniert und in Gl. 4.7b) eingesetzt, so können über die bekannten Zusammenhänge zwischen Schwellspannung und Floating-Gate-Ladung $v_{thH} = Q_{FG}^P / C_{CG}$ und $v_{thL} = Q_{FG}^E / C_{CG}$ Ausdrücke für die Änderung der Schwellspannungen ermittelt werden:

$$\Delta v_{thH} = \frac{\Delta Q_{FG}^P}{C_{CG}} = (1 - (1 - k_D) \frac{x}{t_{ox}}) \frac{\Delta Q_{ox}}{k_D C_{CG}} \quad (4.9)$$

$$\Delta v_{thL} = \frac{\Delta Q_{FG}^E}{C_{CG}} = (1 - k_D) \frac{x}{t_{ox}} \frac{\Delta Q_{ox}}{k_D C_{CG}} \quad (4.10)$$

Mithilfe der Formeln 4.9 und 4.10 lässt sich durch Addition ein linearer Zusammenhang zwischen Oxidladung und Schwellspannungsfenster herstellen.

$$\Delta v_{thW} = \left(\frac{1}{k_D C_{CG}} \right) \Delta Q_{ox} \quad (4.11)$$

Der relative Ladungsschwerpunkt x_R kann leicht über die Gleichungen 4.10 und 4.11 bestimmt werden. Die Kombination von Gl. 4.11 mit 4.10 liefert einen Ausdruck für x_R

in Abhängigkeit von Δv_{thL} und Δv_{thW} .

$$x_R = \frac{\Delta v_{thL}}{\Delta v_{thW}} \frac{1}{1-k_D} \quad (4.12)$$

Neben der Oxidladung und dem Ladungsschwerpunkt ist die injizierte Ladung eine weitere Größe, für die ein Zusammenhang angegeben werden soll. Ähnlich wie die Oxidladung muss die injizierte Ladung direkt aus der Charakteristik der Datenwechselstabilität extrahiert werden, da eine Messung der FN-Ströme nur schwer möglich ist. Da die durch das Oxid geflossene Ladung gleich der Floating-Gate-Ladung ist, kann ein Zusammenhang zwischen Schwellspannung und injizierter Ladung hergestellt werden. Erneut wird der Zusammenhang zwischen Schwellspannung und Floating-Gate-Ladung benutzt.

$$a) \quad v_{thH} = \frac{(Q_{FG}^P = Q_{inj}^P)}{C_{CG}} \quad b) \quad v_{thL} = \frac{(Q_{FG}^E = Q_{inj}^E)}{C_{CG}} \quad (4.13)$$

Werden die Gleichungen 4.13a) und 4.13b) voneinander subtrahiert und die Ladung flächennormiert, erhält man die injizierte Ladung nach einem Programmier- und Löszyklus.

$$Q_{inj}^{P/E} = \frac{2C_{CG}v_{thw}}{A_{inj}} \quad (4.14)$$

Für die Berechnung von $Q_{inj}(z)$ muss über sämtliche P/E -Zyklen summiert werden, was den gesuchten Ausdruck liefert:

$$Q_{inj}(z) = 2 \sum_{m=1}^z Q_{inj}^{P/E}(m) = \frac{2C_{CG}}{A_{inj}} \sum_{m=1}^z v_{thw}(m) \quad (4.15)$$

Mit den Gleichungen 4.11, 4.12 und 4.15 können die Oxidladung, der Ladungsschwerpunkt und die injizierten Ladungen unmittelbar aus den Schwellspannungswerten bestimmt werden. Messungen hierzu sind in Bild 4.21 zu sehen, die den Ladungsschwerpunkt und die Oxidladung in Abhängigkeit von der injizierten Ladung zeigen. Der relative Ladungsschwerpunkt (rechte Achse) wurde mit Gl. 4.12 berechnet und bewegt sich mit zunehmender Zyklenzahl in Richtung Oxidmitte. Der Ladungsschwerpunkt erreicht für $(Q_{inj} \geq 2C/cm^2)$ Positionen nahe der Oxidmitte und ist quasi unabhängig von der injizierten Ladung.

4.4 Diskussion des Ladungseinbaus

Für die Modellierung der Oxidladung stehen verschiedene Modelle zur Verfügung, wie die Diskussion in Kapitel 2 gezeigt hat.

Nachfolgend sollen die Modelle von M. S. Liang und das Potenzgesetz von Papadas et al. zur Beschreibung des Ladungseinbaus angewendet werden. Im Anschluss hieran werden verschiedene Stressformen und die extrahierten Modellparameter miteinander verglichen und diskutiert. Aus Kapitel 2 sind folgende Gesetze bekannt:

4 Untersuchungen zur Oxiddegradation

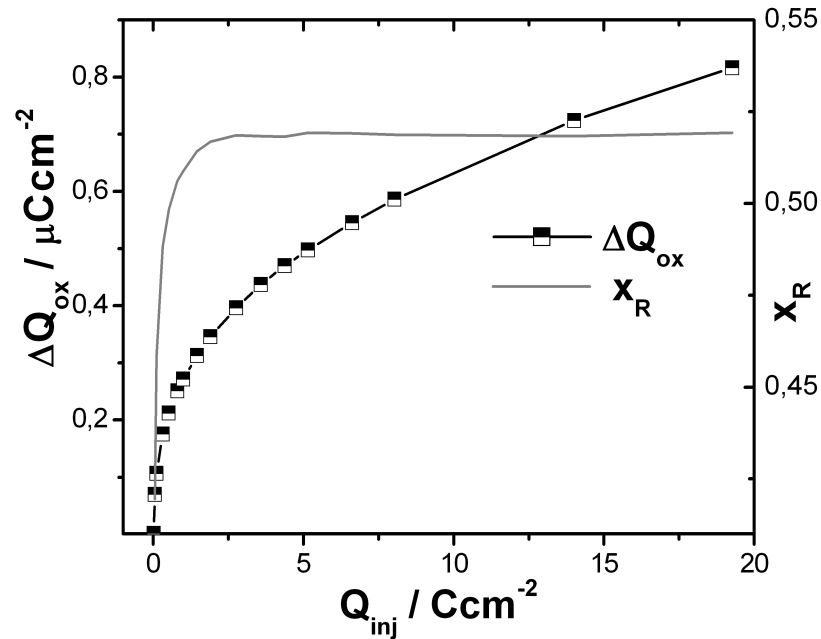


Bild 4.21: Zeigt den Ladungseinbau in einer Speicherzelle. Die Oxidladung (linke Achse) wurde mit Gl.4.11 bestimmt und als Funktion der injizierten Ladung dargestellt.

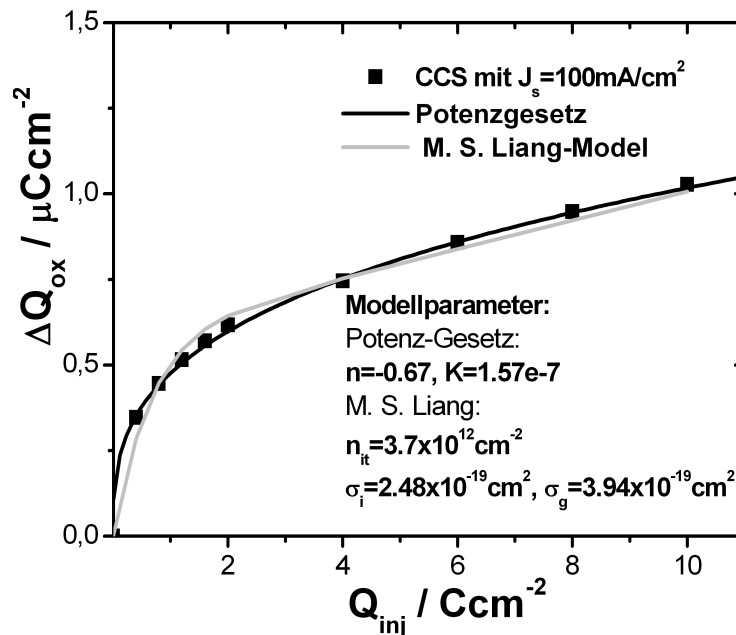


Bild 4.22: Unter CCS ($J_s = +0.1 mA/cm^2$) wurde an einer Injektor-Teststruktur die Oxidladung in Abhängigkeit von der injizierten Ladung gemessen. Das Model von M. S. Liang und die Potenz-Gesetzmäßigkeit von Papadas et al. wurden an die Messergebnisse angepasst.

1) **Papadas-Gesetz:** $\Delta Q_{ox} = \frac{K}{n+1} Q_{inj}^{n+1}$

mit K und n als Fitparameter, wobei K vom Trapping-Yield und n vom Degradationsmechanismus abhängt und zwischen -0.6 und -0.7 liegt.

2) **M. S. Liang:** $\Delta Q_{ox} = eN_{it}(1 - e^{-\sigma_i Q_{inj}}) + \frac{ge^2}{j_s}(Q_{inj} - \frac{1}{\sigma_g}(1 - e^{-\sigma_g Q_{inj}}))$

mit den Wirkungsquerschnitten σ_i und σ_g , der Defekterzeugungsrate g, der Stromdichte j_s und der Trapdichte N_{it} .

In Bild 4.22 wurde an Injektor-Teststrukturen die Oxidladung unter CCS bestimmt und in Abhängigkeit von der injizierten Ladung dargestellt. Für die Anpassung des Papadas-Gesetzes an die Messdaten wurde eine nichtlineare Ausgleichsrechnung durchgeführt. Die gefundenen Modellparameter $K=1.57e-7$ und $n=-0.67$ passen zu den Angaben in [65,66]. Für die Anwendung des Modells von M. S. Liang wurde die Größe ge^2/j_s aus den Messwerten bestimmt. Durch Ableiten von ΔQ_{ox} nach Q_{inj} kann über den Grenzwert für hohe injizierte Ladungen die Größe ge^2/j_s bestimmt werden. Die Wirkungsquerschnitte σ_i , σ_g und die Trapdichte N_{it} konnten im Anschluss hieran durch eine nichtlineare Regression bestimmt werden, siehe Bild 4.23. Es zeigt sich, dass das Potenzgesetz den Ladungseinbau bei den hier betrachteten Teststrukturen besser beschreibt als das Modell von M. S. Liang. Bei der Potenzgesetzmäßigkeit müssen lediglich zwei Modellparameter angepasst werden, wohingegen beim Modell von M. S. Liang vier Parameter σ_i , σ_g , N_{it} und g anzupassen sind. Es zeigt sich weiter, dass die

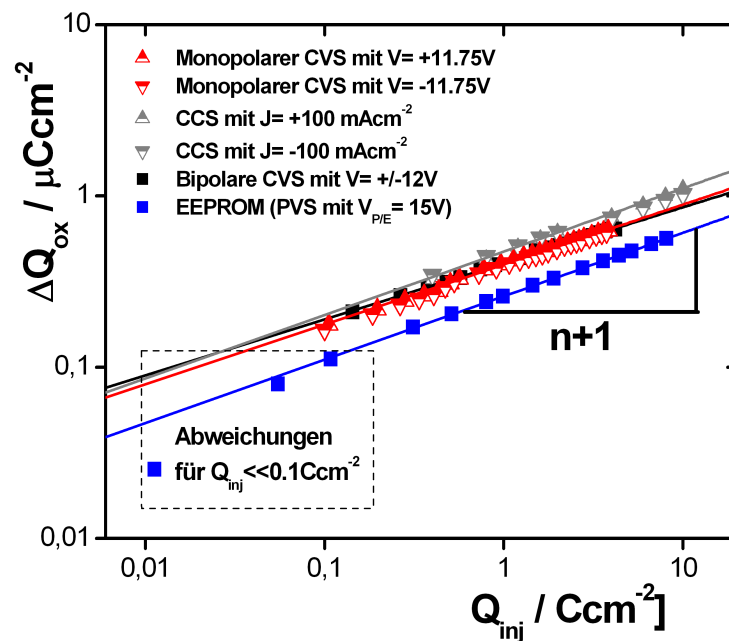


Bild 4.23: Der Zusammenhang $\Delta Q_{ox}(Q_{inj})$ wurde für verschiedene Stressmodi untersucht. Die Potenzgesetzmäßigkeit kann für sämtliche Stressmodi mit hoher Genauigkeit an die Messdaten angepasst werden. Deutliche Unterschiede zeigen sich zwischen statischem (CVS, CCS) und dynamischem Stress (EEPROM).

Potenzgesetzmäßigkeit problemlos auf verschiedene Stressformen anwendbar ist und nicht, wie das Modell von Liang, auf konstanten Stromstress beschränkt ist, siehe Bild 4.23. Der Ladungseinbau wurde hierzu bei verschiedenen Stressmodi gemessen und

4 Untersuchungen zur Oxiddegradation

Stress-Modus	Spannung/Strom	K	n
CVS (Gate)	+11.75V	1.38e-7	-0.63
CVS (Substrat)	-11.75V	1.39e-7	-0.65
CVS (Bipolar)	$\pm 12V$	1.37e-7	-0.67
CCS (Gate)	+100mA	1.50e-7	-0.66
CCS (Substrat)	-100mA	1.70e-7	-0.64
EEPROM	$\pm 15V$	0.96e-7	-0.63

Tabelle 4.1: Zusammenschau der extrahierten Modellparameter K und n für verschiedene Stressmodi und Bauelemente. Der Ladungseinbau der Speicherzellen wurde anhand von 384Bit-Arrays bestimmt. Die Ergebnisse unter CCS und CVS wurden an den Injektor-Teststrukturen verifiziert.

modelliert, siehe Tabelle 4.1. Unabhängig von der Stressform und Injektionsrichtung konnte die Potenzgesetzmäßigkeit an die Messdaten angepasst werden und zeigt lediglich im Fall der Speicherzelle geringe Abweichungen von der Messung.

Die nahezu parallel verlaufenden Geraden (Geradensteigung $n+1$) zeigen, dass sich die n -Parameter nur schwach voneinander unterscheiden. Diese scheinen unabhängig davon zu sein, ob dynamisch oder statisch gestresst wird, Tabelle 4.1. Damit weisen die Unterschiede im Ladungseinbau zwischen EEPROM und Teststruktur auf unterschiedliche K -Parameter bzw. Trapping-Yields hin. Wenn gleiche injizierte Ladungen betrachtet werden, ist die Oxidladung im Tunneloxid der Speicherzelle deutlich geringer als die Oxidladung der Testkapazität. Daher ist der K -Parameter unter statischem Stress deutlich größer als unter dynamischem Stress. Die Unterschiede im Ladungseinbau zwischen dynamischem und statischem Stress werden durch die Frequenzabhängigkeit der Defekterzeugung und Defektbesetzung verursacht. So wurde in [116] gezeigt, dass die Oxiddegradation mit zunehmender Frequenz der Spannungspulse deutlich abnimmt und sich die Datenwechselstabilität verbessert. Aus Messungen zum Stress Induced Leakage Current geht hervor, dass statischer Stress einen deutlich höheren SILC verursacht als dynamischer Stress, was auf eine geringere Oxidschädigung und Defekterzeugung hindeutet [117]. Diese Ergebnisse stehen in Einklang mit Untersuchungen zum Oxiddurchbruch, die ein Ansteigen der „Charge to Breakdown“ mit steigender Pulsfrequenz zeigen [118,119]. Eine Ursache für die Abnahme der Oxiddegradation mit steigender Pulsfrequenz sind die Polaritätswechsel, die mit zunehmender Frequenz zu einer Entleerung von bereits besetzten Defektniveaus führen. Hierbei können grenzflächennahe Defektniveaus durch Tunnelströme beladen und durch Polaritätswechsel wieder entladen werden [120,121]. Die Abnahme der Defekterzeugungsrate mit steigender Pulsfrequenz ist eine weitere Ursache für die Frequenzabhängigkeit der Oxiddegradation. Die Defekterzeugung ist gemäß den Isolatorschädigungsmodellen in Kapitel 2 an die Diffusions- bzw. Driftbewegung von „Spezies“ (Wasserstoffspezies oder Löcher) gebunden, die durch elektrische Felder beeinflusst werden kann.

Bei der anodenseitigen Freisetzung von Löchern „Anode-Hole-Injection“, werden anodenseitig durch Relaxationsprozesse Elektron-Lochpaare erzeugt. Die entstandenen Löcher gelangen durch direktes Tunneln ins Siliziumdioxid und können dort Bindungen aufbrechen. Da der Tunnelprozess reversibel ist, führen die Polaritätswechsel zu einer verlangsamten Bewegung der Löcher in Richtung Kathode. Wenn hohe Frequenzen

vorliegen, sind die Löcher auf den grenzflächennahen Bereich beschränkt. Bei niedrigen Frequenzen können sich die Löcher von der Grenzfläche lösen und Defekte generieren. Analog kann beim „Hydrogen-Release“-Modell argumentiert werden, bei dem die Oxidschädigung auf die Diffusion von Wasserstoffspezies zurückgeführt wird. Die an der Anode freigesetzten Wasserstoffspezies (H^+) können sich bei hohen Frequenzen nicht von den Grenzflächen ablösen, was die Oxidschädigung verlangsamt [122,123].

4.5 Untersuchungen zu Modellparametern

Im Anschluss an die Untersuchungen zum Ladungseinbau wird die Abhängigkeit der n - und K -Parameter von der Injektionsrichtung, der Stromdichte und von der Temperatur näher untersucht. Des Weiteren wird der Ladungseinbau der Speicherzellen mit dem der Testkapazitäten verglichen und diskutiert.

4.5.1 Richtungsabhängigkeit

Wie aus Tabelle 4.1 hervorgeht, ist der Ladungseinbau der hier betrachteten Testkapazitäten nur schwach richtungsabhängig, sodass die Spannungspolarität und die Stromrichtung einen geringen Einfluss auf die Modellparameter ausüben. Die Modellparameter weisen auf eine störungsfreie Überlagerung aus gate- und substratseitigem Ladungseinbau hin, wodurch sich dieser unter monopolarer und bipolarer Stress durch eine einheitliche Fitfunktion modellieren lässt.

Eine geringe Richtungsabhängigkeit des Ladungseinbaues findet sich z. B. auch in den Untersuchungen von C. Busseret et al. [124]. Es gibt allerdings auch Untersuchungen [125,126], die eine deutliche Richtungsabhängigkeit des Ladungseinbaues bzw. der n - und K -Parameter nachweisen.

Ferner kann ein einfacher Zusammenhang zwischen dem Ladungseinbau unter bipolarer und monopolarer Stress hergestellt werden, wenn die Ladungsschwerpunkte und Oxidladungen unter monopolarer Stress für beide Injektionsrichtungen bekannt sind [124]. Mithilfe des Schwerpunktsatzes kann hieraus die Oxidladung und der Ladungsschwerpunkte für den bipolaren Fall berechnet werden,

$$x^{Bi} = \frac{\Delta Q_{ox}^G x_G + \Delta Q_{ox}^S x_S}{\Delta Q_{ox}^G + \Delta Q_{ox}^S} \quad \text{mit} \quad \Delta Q_{ox}^{Bi} = \Delta Q_{ox}^G + \Delta Q_{ox}^S$$

wobei x_G und x_S die gate- bzw. substratseitigen Ladungsschwerpunkte sind. ΔQ_{ox}^G und ΔQ_{ox}^S bezeichnen die jeweiligen Oxidladungen.

4.5.2 Stromdichte- und Temperaturabhängigkeit

Unter Verwendung von CCS wurde $\Delta Q_{ox}(Q_{inj})$ für verschiedene Stromdichten gemessen und in Bild 4.24 dargestellt.

Im ersten Teilabschnitt lagen die verwendeten Stromdichten zwischen $1 \text{ mAc}m^{-2}$ und $50 \text{ mAc}m^{-2}$. Deutlich ist eine schwache Zunahme der eingebauten Oxidladung mit steigender Stromdichte zu erkennen. Im zweiten Teilabschnitt wurden Stromdichten von 0.1 , 0.2 und $0.5 \text{ Ac}m^{-2}$ angewendet. Ein Oxiddurchbruch konnte bei ca. $10 \text{ Cc}m^{-2}$

4 Untersuchungen zur Oxiddegradation

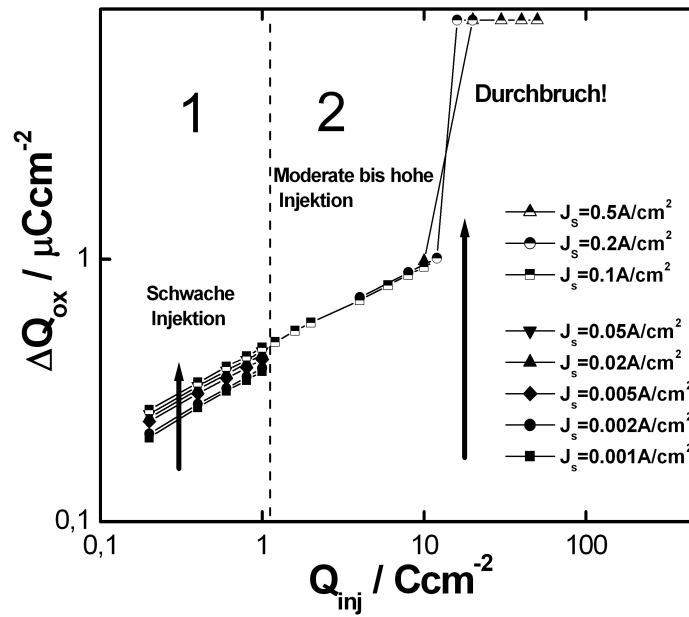


Bild 4.24: Darstellung der Oxidladung als Funktion der injizierten Ladung für verschiedene Stromdichten. Im ersten Teilabschnitt wurde $\Delta Q_{ox}(Q_{inj})$ für Stromdichten zwischen $j=1 \text{ mA/cm}^2$ und 50 mA/cm^2 untersucht. In Teilabschnitt 2 lagen die verwendeten Stromdichten bei $0.1, 0.2$ und 0.5 Acm^{-2} .

beobachtet werden. Die aus den Messdaten extrahierten n - und K -Parameter sind in Bild 4.25 in Abhängigkeit von der Stromdichte dargestellt, wobei die Parameter für beide Injektionsrichtungen bestimmt wurden. Sowohl der n -Parameter als auch der K -Parameter sind nur schwach richtungsabhängig. Weiterhin zeigt sich eine schwach ausgeprägte Stromdichteabhängigkeit des n -Parameters, wohingegen der K -Parameter mit steigender Stromdichte deutlich zunimmt. In [110] wird für den Zusammenhang zwischen K -Parameter und Stromdichte folgende Potenzgesetzmäßigkeit angegeben.

$$K(j_s) \propto j_s^m \text{ mit } m \text{ als Fitparameter}$$

Die Ergebnisse zeigen, dass die Zunahme der Oxidladung durch die Änderung des K -Parameters bzw. des Trapping-Yields verursacht wird und deutlich von der Stromdichte abhängt. Eine weitere Untersuchung wurde hinsichtlich der Temperaturabhängigkeit des Ladungseinbaues durchgeführt. Für verschiedene Temperaturen wurde $\Delta Q_{ox}(Q_{inj})$ gemessen und eine Zunahme der Oxidladung mit steigender Temperatur festgestellt, Bild 4.26. Der Zusammenhang zwischen Oxidladung und injizierter Ladung kann auch im Bereich hoher Temperaturen mit dem Potenzgesetz von Papadas modelliert werden. Der nahezu parallele Verlauf der Geraden deutet auf eine schwache Temperaturabhängigkeit des n -Parameters hin.

Die extrahierten n - und K -Parameter sind in Bild 4.27 dargestellt. Sie sind über der reziproken thermischen Energie aufgetragen. Der n -Parameter (rechte Achse) schwankt zwischen -0.68 und -0.69 und erweist sich als quasi temperaturunabhängig, wobei der K -Parameter signifikant mit steigender Temperatur zunimmt. Der Arrhenius-Plot zeigt, dass der K -Parameter thermisch aktivierbar ist und eine Aktivierungsenergie von $\epsilon_a=0.042\text{eV}$ besitzt. Der Befund eines quasi temperaturunabhängigen n -Parameters und

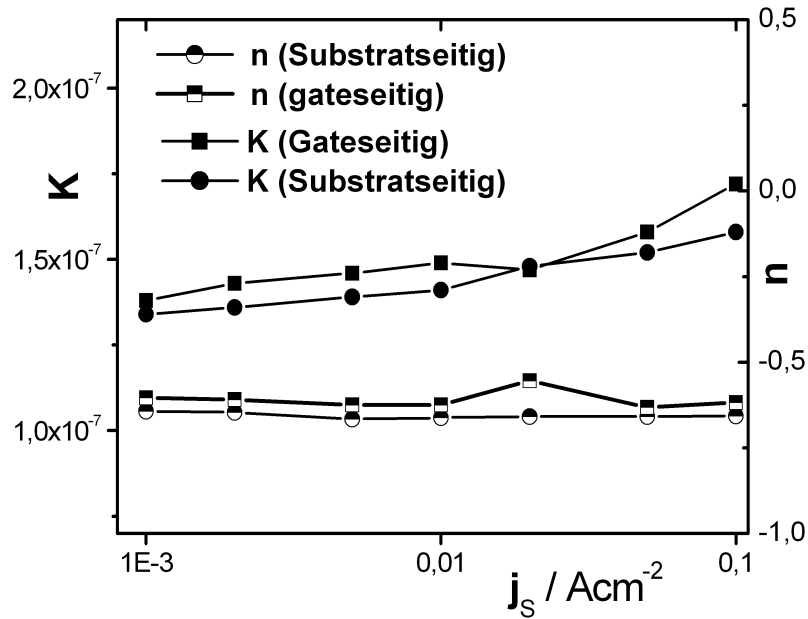


Bild 4.25: Der K-Parameter (linke Achse) und der n-Parameter (rechte Achse) für gate- und substratseitige Injektion sind in Abhängigkeit von der verwendeten Stromdichte dargestellt.

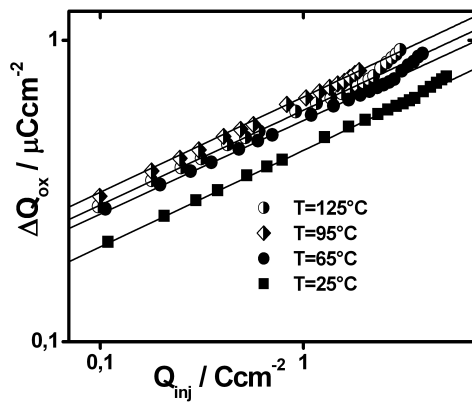


Bild 4.26: Die Oxidladung in Abhängigkeit von Q_{inj} wurde für verschiedene Temperaturen gemessen und über ein Potenzgesetz modelliert.

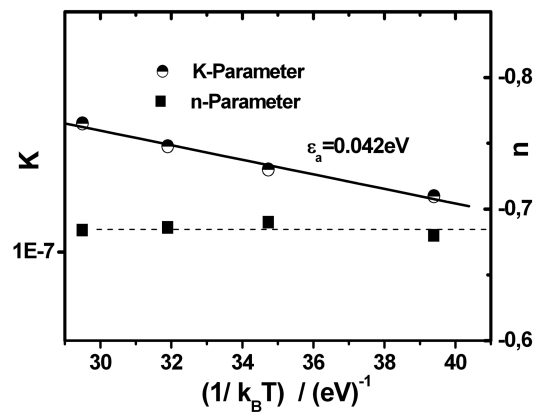


Bild 4.27: Der K-Parameter (linke Achse) und der n-Parameter (rechte Achse) wurden halblogarithmisch über der reziproken thermischen Energie aufgetragen.

4 Untersuchungen zur Oxiddegradation

die für den K -Parameter ermittelte Aktivierungsenergie stimmen gut mit den Untersuchungen in [113] überein.

4.5.3 Korrelation zwischen Speicherzelle und Teststruktur

Der Ladungseinbau im Tunneloxid der Speicherzellen soll im Folgenden mit dem der Injektor-Teststrukturen verglichen werden, wobei die Größen der Teststruktur (Speicherzelle) mit einem T bzw. (E) gekennzeichnet werden. In Bild 4.28 sind die Resultate für beide Bauteile dargestellt. Der Ladungseinbau der Speicherzelle wurde für ver-

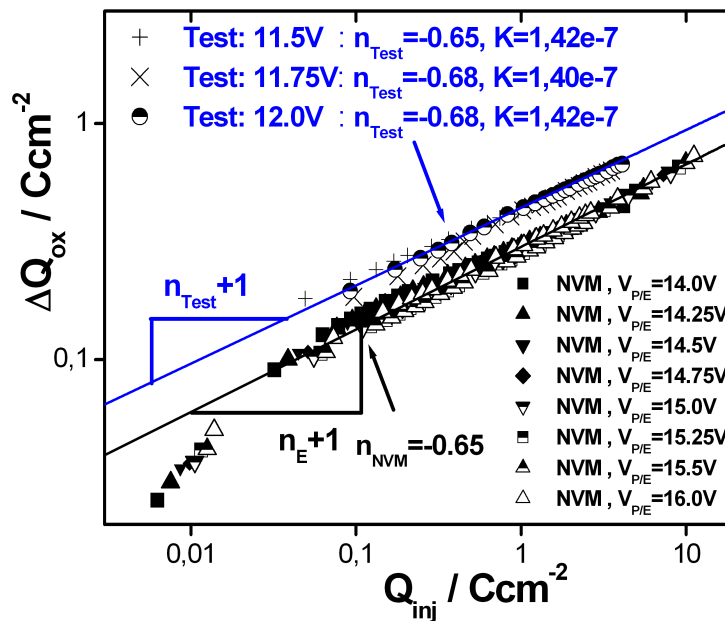


Bild 4.28: Die Oxidladung $\Delta Q_{\text{ox}}(Q_{\text{inj}})$ wurde an EEPROM und Injektor-Teststrukturen gemessen. Die Oxide im EEPROM degradieren deutlich langsamer als die Teststrukturen unter bipolarem CVS.

schiedene Programmier- und Löschspannungen untersucht und erwies sich als schwach spannungsabhängig. Die leichte Zunahme des Ladungseinbaues mit zunehmender P/E -Spannung kann deutlich an den extrahierten K -Parametern erkannt werden, die in Tabelle 4.2 zusammengestellt sind. Im Vergleich zur Speicherzelle fällt die Degradation der Teststruktur unter CVS höher aus, was sich in den unterschiedlichen K -Parametern zeigt. Die n -Parameter der Teststrukturen liegen zwischen -0.67 und -0.7 und weichen weniger als 10 % von den n -Werten der Speicherzelle ab (siehe Tabelle 4.2 und Bild 4.28)

Ähnliche Resultate liefert der Vergleich zwischen Injektor-Teststruktur und Speicherzellen, die mit unterschiedlichen Pulslängen programmiert und gelöscht wurden, Bild 4.29. Mit abnehmender Pulslänge kann ein leichter Rückgang des Ladungseinbaues bzw. der Oxidladung beobachtet werden. Die n -Parameter der Speicherzellen unterscheiden sich geringfügig von den Werten der Teststrukturen (siehe Tabelle 4.2 und Bild 4.29). Die geringe Abweichung zwischen den n -Parametern impliziert, dass das

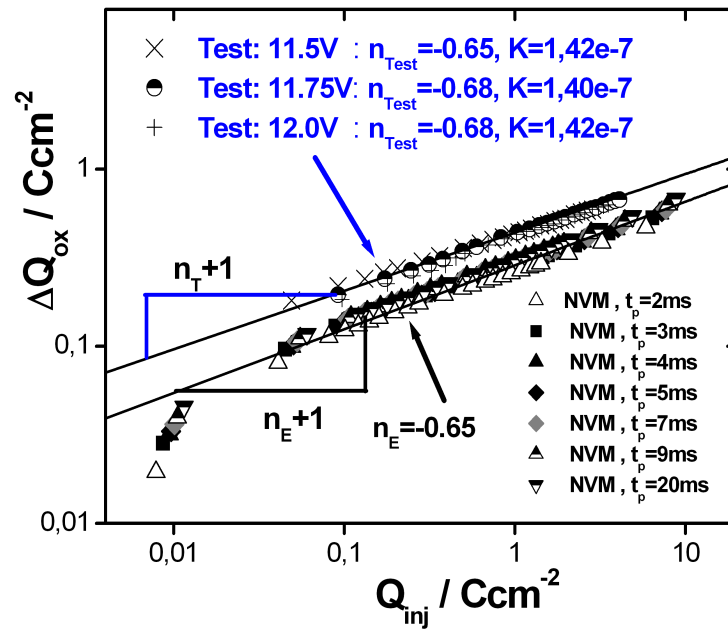


Bild 4.29: Die Oxidladung $\Delta Q_{ox}(Q_{inj})$ wurde an EEPROM und Injektor-Teststrukturen bestimmt. Auch hier degradieren die Tunneloxide der EEPROM schneller als die Oxide der Injektor-Teststrukturen.

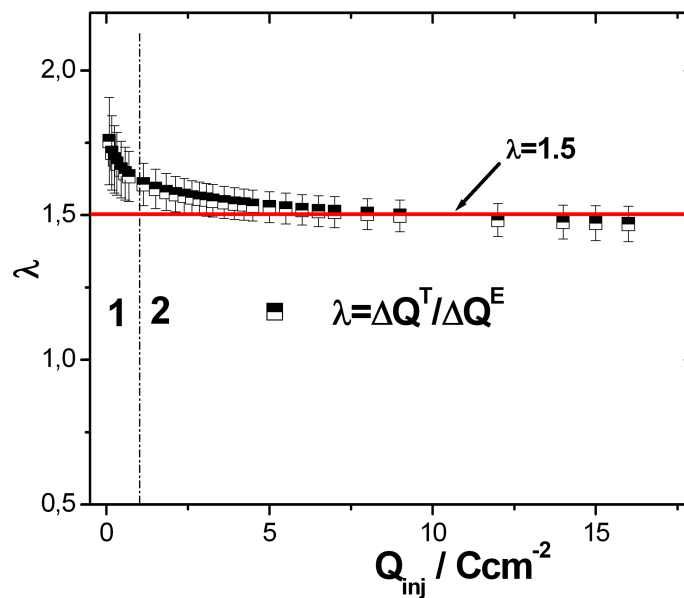


Bild 4.30: Darstellung des λ -Parameters als Funktion der Ladungsinjektion. Für schwache Ladungsinjektionen (Bereich 1) liegt λ zwischen 1.6 und 1.7 und ist noch deutlich von der injizierten Ladung abhängig. Bei höheren Ladungsinjektionen (Bereich 2) liegt Lambda zwischen 1.5 und 1.6 und hängt schwach von der Ladungsinjektion ab.

4 Untersuchungen zur Oxiddegradation

Bild 4.29	$V_{P/E}$	t_p	K	n
	15.0V	2ms	8.53e-8	-0.67
	15.0V	3ms	9.39e-8	-0.67
	15.0V	4ms	9.71e-8	-0.66
	15.0V	5ms	9.93e-8	-0.66
	15.0V	7ms	1.03e-7	-0.66
	15.5V	9ms	1.05e-7	-0.65
	15.0V	20ms	1.09e-7	-0.64
Bild 4.28	14.0V	9ms	8.46e-8	-0.71
	14.25V	9ms	9.25e-8	-0.69
	14.5V	9ms	9.96e-8	-0.68
	14.75V	9ms	1.03e-7	-0.66
	15.0V	9ms	1.03e-7	-0.64
	15.25V	9ms	1.05e-7	-0.63
	15.5V	9ms	1.05e-7	-0.62
	16.0V	9ms	1.07e-7	-0.64

Tabelle 4.2: Zusammenschau der extrahierten Modellparameter für die Speicherzellen aus Bild 4.28 und 4.29

Verhältnis der Oxidladungen ΔQ_{ox}^T und ΔQ_{ox}^E nur schwach von den injizierten Ladungen abhängt. Um dies näher zu untersuchen, wurde für 10 Wafer das Verhältnis aus beiden Oxidladungen (λ) bestimmt. Die Mittelwerte sowie die Standardabweichung wurden über der injizierten Ladung aufgetragen, siehe Bild 4.30. Für injizierte Ladungen kleiner 1 Ccm^{-2} (Bereich 1) liegt λ zwischen 1.6 und 1.8 und ist noch deutlich von Q_{inj} abhängig.

Für höhere Ladungsinjektionen (Bereich 2) zeigt λ eine schwache Abhängigkeit von Q_{inj} und es gilt näherungsweise $\lambda \approx 1.5$. In diesem Bereich ist die Oxidladung im Tunneloxid der Speicherzelle direkt proportional zur Oxidladung der Teststruktur. Die geringe Standardabweichung ($\leq 5 \%$) weist ferner auf die gute Reproduzierbarkeit der Messergebnisse hin.

4.6 Entwicklung eines schnellen EEPROM-Testverfahrens

Die Ergebnisse zum Ladungseinbau eröffnen die Möglichkeit, ein WLR (Wafer Level Reliability)-Testverfahren zu konzipieren, mit dem sich die Oxidqualität und die Zykelfestigkeit von Speicherzellen bestimmen lassen. Das Testverfahren basiert auf einer beschleunigten Messung des Ladungseinbaues anhand von einfachen Testkapazitäten. Hierüber kann der Ladungseinbau der Speicherzelle berechnet bzw. extrapoliert werden. Das Testkonzept und seine Implementierung soll nachfolgend vorgestellt werden.

4.6.1 Anforderungsprofil an beschleunigte Testverfahren

In einer Qualifikation durchläuft ein Automotivprodukt eine Vielzahl von Tests, um die Produktmerkmale zu überprüfen. Die Testverfahren sollen das Ausfallverhalten der Produkte untersuchen und primär den dritten Teil der „Badewannenkurve“ erfassen [127], in dem Verschleißausfälle (Materialermüdung) dominieren.

Die Zielsetzung ist es hierbei, Qualitätsmerkmale wie sie z. B. in Teststandards gefordert werden, sicherzustellen. Eine weitere Aufgabe von Testverfahren besteht im Auffinden von Material- oder Layoutfehlern, um neue Produkte während ihrer Entwicklungsphase optimieren zu können. Problematisch ist hierbei, dass viele Ausfallmechanismen statistischer Natur sind und der damit verbundene Zeitaufwand nur mit zeitraffenden, beschleunigten Testverfahren gering gehalten werden kann, um zeitnah zur Produktentwicklung auf Fehlerquellen reagieren zu können. Bei der Entwicklung von Testverfahren sind daher folgende Aspekte von besonderem Interesse:

- Geringe Testzeit bzw. hohes Beschleunigungsvermögen
- Einfache Testidee und Implementierung
- Gleiche Fehlermechanismen unter Feld- und Testbeanspruchung
- Korrelation zwischen Testbauteil und Produkt

Neben einer möglichst geringen Testzeit muss sichergestellt werden, dass das Testverfahren die gleichen Fehlermechanismen adressiert wie die Feldbeanspruchung in der Applikation. Eine Überbeanspruchung des Bauteiles unter Testbedingungen könnte Fehlermechanismen initiieren, die am Bauteil unter Feldbeanspruchung bedeutungslos sind. Somit wären über das Testverfahren keine Rückschlüsse auf das Ausfallverhalten des Bauteiles unter Feldbelastung möglich. Wird ein Testbauteil oder eine Teststruktur verwendet, so muss zusätzlich die Vergleichbarkeit zwischen beiden Bauteilen sichergestellt sein.

4.6.2 Evaluierung des Ausfallverhaltens über Indikatorgrößen

Die Degradation eines Bauteiles zieht die Veränderung mehrerer physikalischer Größen nach sich, die charakteristisch für die Funktionsweise des Bauteiles sind. Das Veränderungsverhalten dieser Größen kann direkt oder indirekt mit einem Ausfallmechanismus korrelieren, sodass hierüber auf den Zeitpunkt des Ausfalles geschlossen werden kann.

4 Untersuchungen zur Oxiddegradation

Eine solche physikalische Größe soll nachfolgend als Stressindikator (I) bezeichnet werden und wird in der Regel in Abhängigkeit von einer akkumulierten Stressgröße dargestellt, siehe Bild 4.31. So ist z. B. der ohmsche Widerstand eines Bond-Drahtes

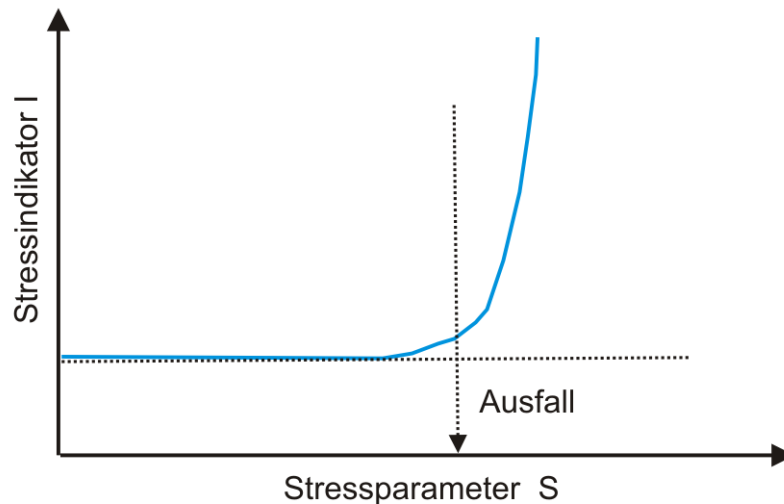


Bild 4.31: Schematischer Verlauf eines Stressindikators I als Funktion eines Stressparameters. Ein Bauelementausfall liegt vor, wenn die Indikatorgröße einen kritischen Wert überschreitet.

ein geeigneter Stressindikator, um den Ausfall eines Bauteiles durch eine fehlerhafte Verbindungsstelle zu untersuchen [127]. Die Bond-Draht-Verbindungsstellen werden durch thermische Wechselbelastung geschwächt, sodass sich der ohmsche Widerstand nachhaltig verändert. Die Rissbildung innerhalb des Drahtmaterials oder ein vollständiger Drahtbruch führen schließlich zu einem abrupten Anstieg des Widerstandes, der den Ausfallzeitpunkt der Verbindungsstelle bzw. des Bauteils markiert. Da jedes Ausfallbild durch mindestens einen Ausfallmechanismus hervorgerufen wird, muss zu jedem Ausfallbild ein geeigneter Stressindikator gefunden werden.

Neben der Bestimmung von geeigneten Stressindikatoren muss die Frage beantwortet werden, wie die Degradation des Bauteiles zu beschleunigen ist, um Testzeit und Testkosten zu senken, ohne dass hierbei Fehlermechanismen auftreten, die unter normaler Feldbelastung nicht relevant sind. Um Degradationsprozesse zu beschleunigen, gibt es folgende Möglichkeiten:

- Erhöhung der Umgebungstemperatur
- Erhöhung der Stressbelastung durch die Erhöhung von Strom- und Spannungsamplituden sowie der Einwirkdauer
- Höhere Stressbelastung durch zeitliche Raffung von Wechselbelastungen
- Kombination verschiedener Stressformen.

Die Erhöhung der Umgebungstemperatur wird eingesetzt, um thermisch aktivierbare Fehlermechanismen zu untersuchen und zu beschleunigen. Hierzu zählen Ausfallmechanismen, die auf Migrations- und Diffusionsvorgänge, Phasenbildung und Materialausgasungen zurückzuführen sind. Die Umgebungstemperatur kann allerdings nicht

beliebig weit gesteigert werden, da bei zu hohen Temperaturen neben den eigentlich zu beschleunigenden Fehlermechanismen neue Fehlermechanismen hinzutreten. Im Zusammenhang mit der erhöhten thermischen Belastung sei noch die zeitliche Raffung einer Temperaturänderung zu erwähnen, wie sie z. B. täglich in Automobilen stattfindet. Im TST (Temperatur-Schock-Test) [128] werden die Bauelemente zwischen zwei Kammersystemen mit unterschiedlichen Temperaturen hin- und hergefahren, um Temperaturänderungen unter Feldbelastung zeitlich gerafft abzubilden.

Eine weitere Möglichkeit Degradationsprozesse zu beschleunigen, besteht in der Erhöhung von Strom und Spannungsgrößen. In der Verbindungstechnik wird beim sogenannten Power-Cycling-Test [128] durch eine überdurchschnittlich hohe Strombelastung thermomechanischer Stress erzeugt, der die Verbindungstechnik beschleunigt altern lässt und z.B. die vorzeitige Ablösung einer Drahtverbindung initiiert. Auch bei Speicherzellen kann durch Erhöhung der Programmier- und Löschspannung die Degradation des Tunneloxides bzw. des Schwellspannungsfensters (Datenwechselstabilität) beschleunigt werden.

Die Kombination verschiedener Stressformen, wie z. B. einer Temperatur- und Spannungsbelastung, wäre eine weitere Möglichkeit, Alterungsprozesse zu beschleunigen. Problematisch ist hierbei die eindeutige Zuordnung zwischen Ausfallbild und Ausfallursache. Insbesondere bei eingehäuteten Chips ist die eindeutige Zuordnung zwischen Fehlerbild und Fehlermechanismus schwierig, da in solchen Systemen meistens ein komplexes Zusammenspiel verschiedener Fehlermechanismen stattfindet. Es gibt hierzu Testverfahren, die gezielt das Auftreten verschiedener Stressformen nutzen und durch die Verstärkung einer Stressform den Alterungsprozess beschleunigen bzw. den Ausfall vorzeitig herbeiführen. In [129] wird z. B. ein Testverfahren zur Lotstellenermüdung vorgestellt, das thermische Wechselbelastung mit einer mechanischen Zusatzbelastung kombiniert, um die in den Lotschichten auftretenden thermomechanischen Spannungen nachzubilden und zu verstärken.

Die Lotstellenermüdung ist auf die Ausbildung von Rissen innerhalb der Lotschicht zurückzuführen, die zur Delamination der Lotschicht und schließlich zum Ausfall des Bauteiles führen. Wird z. B. die Platine mit einem Biegemoment beaufschlagt, so werden die Lotschichten einer mechanischen Zusatzbelastung ausgesetzt. In Kombination mit einer Temperaturbelastung kann so die Rissbildung innerhalb der Lotschicht beschleunigt werden. Simulationen hierzu zeigen [129], dass bei einer geeigneten Auswahl der mechanischen Zusatzbelastung (Biegemoment) die gleichen Ausfallmechanismen aktiviert werden, wie im Fall der reinen Temperaturbelastung.

4.6.3 WLR-Test zur Bestimmung der EEPROM-Datenwechselstabilität

Die allgemeinen Bemerkungen zum Anforderungsprofil eines Testverfahrens haben gezeigt, dass für jeden Ausfallmechanismus ein geeigneter Stressindikator zu spezifizieren ist. Hieran schließt sich die Frage an, wie dieser gemessen und wie das Testverfahren gegenüber der Feldbelastung beschleunigt werden kann.

Soll ein Testverfahren auf Wafer-Ebene für die Datenwechselstabilität und Oxidqualität von EEPROM entwickelt werden, muss zuerst ein Stressindikator und ein geeignetes Messverfahren zu seiner Bestimmung gefunden werden. Wie aus den Untersuchungen

4 Untersuchungen zur Oxiddegradation

zum Ladungseinbau hervorging, ist die Oxidladung eindeutig mit der Degradation des Tunneloxides korreliert und dadurch ein geeigneter Stressindikator für die Oxiddegradation im EEPROM. Die Oxidladung wurde einerseits über die DiMaria-Methode und das Papadas-Verfahren ermittelt und in Abhängigkeit von der akkumulierten injizierten Ladung dargestellt. Basierend auf den Messungen von $\Delta Q_{ox}(Q_{inj})$ soll nachfolgend ein rechnerisches Verfahren entwickelt werden, mit dem die Datenwechselstabilität der EEPROM über den Ladungseinbau der Testkapazität simuliert und bewertet werden kann. Das Testprinzip ist schematisch in Bild 4.32 dargestellt und umfasst folgende Punkte:

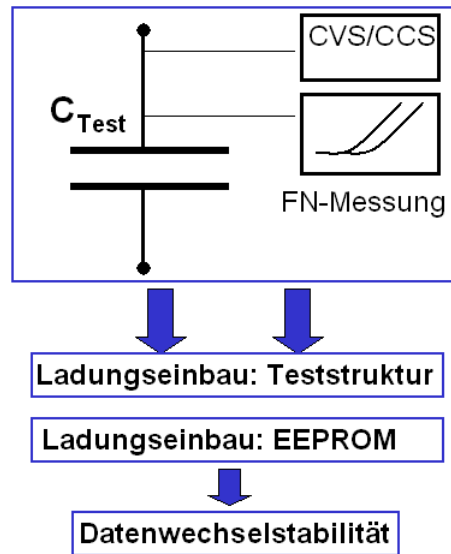


Bild 4.32: Schematische Darstellung des WLR-Tests zur Bestimmung der Datenwechselstabilität und Oxidqualität in EEPROM Bauteilen.

1. Unter Verwendung von statischem Stress werden die Tunneloxidkapazitäten gealtert und der Ladungseinbau $\Delta Q_{ox}^T(Q_{inj})$ mithilfe der DiMaria-Methode gemessen.
2. Mithilfe eines analytischen Zusammenhanges kann der Ladungseinbau der Speicher über den Ladungseinbau der Teststrukturen bestimmt werden.
3. Hierüber erfolgt die Simulation der Datenwechselstabilität und deren Bewertung.

Die Messungen des Ladungseinbaues haben einerseits gezeigt, dass der Ladungseinbau im EEPROM-Bauteil deutlich langsamer abläuft als bei den Tunneloxidkapazitäten unter Verwendung von statischem Stress (CCS und CVS). Ursache hierfür waren die in Kapitel 4.4 erläuterten Unterschiede zwischen PVS und CVS/CCS. Daher kann durch Verwendung von statischem Stress die Testzeit gegenüber einer Messung an Speicherzellen signifikant gesenkt werden. Die Vergleichbarkeit zwischen Speicherzelle und Testkapazität setzt ferner voraus, dass in beiden Fällen bipolarer Stress verwendet werden muss [130-132]. Würden hingegen die Testkapazitäten unter monopolarer Stress degradieren und die Speicherzellen unter bidirektionalem Stress, so wären die Ergebnisse aufgrund einer möglichen Richtungsabhängigkeit des Ladungseinbaues nicht miteinander vergleichbar. Die im Rahmen der Modellierung eingeführten Modellparameter

n und K zeigen, dass der Ladungseinbau unter CVS etwa 1.4- bis 1.6-mal schneller abläuft als unter PVS. Dies folgt deutlich aus dem Verhältnis der K -Parameter, die als Maß für den Trapping-Yield zu interpretieren sind. Die geringe Abhängigkeit der n -Parameter vom Stressmodus zeigt ferner, dass der Degradationsmechanismus unabhängig vom Stressmodus ist, also davon, ob dynamischer oder statischer Stress verwendet wird. Dies impliziert einen Fehlermechanismus, der ebenfalls unabhängig vom Stressmodus ist. Somit werden im Tunneloxid unter Test- bzw. Feldbelastung die gleichen Fehlermechanismen angesprochen.

Nach der hier gegebenen Übersicht über das Testkonzept soll nachfolgend ein Zusammenhang zwischen dem Ladungseinbau bzw. den Stressindikatoren beider Bauteile hergeleitet werden. Im Anschluss hieran wird ein rechnerisches Verfahren entwickelt, um die Datenwechselstabilität und Oxidqualität beurteilen zu können.

4.6.4 WLR-Simulationsverfahren

Die Messungen in Kap. 4.5.3 haben gezeigt, dass die Oxidladung der Speicherzelle direkt proportional zur Oxidladung der Teststruktur ist, wenn die betrachteten injizierten Ladungen gleich sind.

Die Ursache hierfür ist die geringe Abhängigkeit des n -Parameters vom verwendeten Stressmodus, sodass n_T und n_E nur geringfügig differieren und näherungsweise gleich sind $n^T \approx n^E$. Wird für die Modellierung der Speicherzelle und der Testkapazität das Papadas-Gesetz angewandt, so folgt

$$\text{Teststruktur} \quad \Delta Q_{ox}^T = \frac{K_T}{n_T + 1} (Q_{inj})^{n_T + 1} \quad \text{EEPROM} \quad \Delta Q_{ox}^E = \frac{K_E}{n_E + 1} (Q_{inj})^{n_E + 1} \quad (4.16)$$

Betrachtet man für beide Bauteile den gleichen Ladungsdurchfluss ($Q_{inj}^T = Q_{inj}^E$), oder die gleichen Oxidladungen ($\Delta Q_{ox}^T = \Delta Q_{ox}^E$) so findet man nach Division der beiden Gleichungen folgende Relationen:

$$a) \quad \frac{\Delta Q_{ox}^T}{\Delta Q_{ox}^E} \approx \frac{K^T}{K^E} = \lambda \quad b) \quad \frac{Q_{inj}^T}{Q_{inj}^E} \approx \left(\frac{1}{\lambda}\right)^{1/n+1} \quad (4.17)$$

Bei einer gegebenen injizierten Ladung kann hierüber die Oxidladung der Teststruktur in die korrespondierende Oxidladung der Speicherzelle umgerechnet werden.

Für die Simulation der Datenwechselstabilität ist neben den Oxidladungen ein Zusammenhang zwischen Zykelzahl und injizierter Ladung erforderlich, um eine gemessene Ladungsinjektion in die dazu äquivalente Zykelzahl umzurechnen. Aus der Grundgleichung für die injizierte Ladung Gl. 4.15 kann durch Linearisierung folgender Ausdruck gefunden werden:

$$Q_{inj}(z) = \frac{2C_{CG}}{A_{inj}} \sum_{m=1}^z v_{thw}(m) = \frac{2C_{CG}}{A_{inj}} \sum_{m=1}^z \left(v_{thw0} + \frac{dv_{thw}(m)}{dm} m + \dots \right) \quad (4.18)$$

$$Q_{inj} \approx \frac{2C_{CG} v_{thw0}}{A_{inj}} z \quad (4.19)$$

4 Untersuchungen zur Oxiddegradation

Dieses „lineare“ Modell basiert auf der Annahme, dass im Bereich schwacher bis moderater Degradationen das Schwellspannungsfenster näherungsweise konstant ist. Da das Schwellspannungsfenster für hohe Degradationen deutlich abnimmt, wird durch Gl. 4.19 der Ladungsfluss überschätzt, sodass diese Näherung und die Grenzen ihrer Anwendbarkeit noch genauer betrachtet werden müssen.

Zusammenfassend ergeben sich für die Zyklenzahlen und die korrespondierenden Oxidladungen der Speicherzelle folgende Gleichungen:

(1) Wenn $Q_{inj}^T = Q_{inj}^E$

$$a) \quad z = \frac{Q_{inj}^T A_{inj}}{2C_{CG} v_{thw0}} \quad b) \quad \Delta Q_{ox}^E = \frac{1}{\lambda} \Delta Q_{ox}^T \quad (4.20)$$

Für den Fall gleicher Oxidladungen muss für die Berechnung der Zyklenzahl Gl. 4.19 mit Gl. 4.17 b) kombiniert werden. Es folgt hieraus:

(2) Wenn $\Delta Q_{ox}^T = \Delta Q_{ox}^E$:

$$a) \quad z = \frac{Q_{inj}^T \left(\frac{K^T}{K^E}\right)^{\frac{1}{n+1}} A_{inj}}{2C_{CG} v_{thw0}} \quad b) \quad \Delta Q_{ox}^E \quad (4.21)$$

Über die Oxidladung der Speicherzelle kann auf direktem Weg die Schwellspannungsänderung bestimmt werden, was über die in Kapitel 4.3 hergeleiteten Gleichungen 4.9 und 4.10 geschieht. Ausgangspunkt sind die an der Teststruktur ermittelten Größen Q_{inj}^T und ΔQ_{ox}^T . Aus der injizierten Ladung kann über Gl. 4.20 a) die Zyklenzahl berechnet werden. Wird Gl. 4.20 b) in die Gleichungen 4.9 und 4.10 eingesetzt, so folgen Gleichungen für die Änderung der Schwellspannung im programmierten und gelöschten Zustand (Grundgleichungen des WLR-Testes).

$$\Delta v_{thH}(\Delta Q_{ox}^T) = (1 - (1 - k_D)) \frac{x}{t_{ox}} \frac{\left(\frac{1}{\lambda} \Delta Q_{ox}^T = \Delta Q_{ox}^E\right)}{k_D C_{CG}} \quad \text{mit} \quad z = \frac{Q_{inj}^T A_{inj}}{2C_{CG} v_{thw0}} \quad (4.22)$$

$$\Delta v_{thL}(\Delta Q_{ox}^T) = (1 - k_D) \frac{x}{t_{ox}} \frac{\left(\frac{1}{\lambda} \Delta Q_{ox}^T = \Delta Q_{ox}^E\right)}{k_D C_{CG}} \quad \text{mit} \quad z = \frac{Q_{inj}^T A_{inj}}{2C_{CG} v_{thw0}} \quad (4.23)$$

Das Schwellspannungsfenster v_{thw0} wird mit dem in Kapitel 3 entwickelten Modell bestimmt, indem die Schwellspannungen im programmierten und (gelöschten) Zustand v_{thH0} und (v_{thL0}) simuliert werden. Hierbei bezeichnen v_{thH0} und v_{thL0} die Schwellspannungen nach dem ersten P/E-Zyklus ($\Delta Q_{ox} \approx 0$). Mit diesen Werten können die Schwellspannungen als Funktion der Zyklenzahl berechnet werden.

4.6.5 Simulationsergebnisse

Nachfolgend wird das WLR-Verfahren zur Simulation der Datenwechselstabilität angewandt. Ausgangspunkt sind Stressmessungen an Testkapazitäten, die sich im gleichen Belichtungsfeld wie die 384Bit-EEPROM-Arrays befinden. Der Ladungseinbau erfolgt unter bipolarem CVS und liefert die Inputdaten für das WLR-Verfahren. Die darüber simulierten Schwellspannungsverläufe werden zusammen mit den mittleren

Schwellsprungwerten des 384Bit-Arrays in Abhängigkeit von der P/E-Zykelzahl dargestellt. In Bild 4.33 sind Simulationen mit verschiedenen λ -Werte zu sehen, die unterschiedlich starke Korrelationen zwischen Speicherzelle und Testkapazität zeigen. Die Ergebnisse zum Ladungseinbau haben gezeigt, dass im Bereich schwacher Degradation die λ -Werte größer 1.7 sind. Im Bereich moderater bis hoher Ladungsinjektion ist λ schwach von der injizierten Ladung abhängig und liegt zwischen 1.5 und 1.6. Das WLR-Verfahren liefert für $\lambda \geq 1.7$ Schwellsprungsverläufe die unterhalb (gelöscht) bzw. oberhalb (programmiert) der Messergebnisse liegen. Die Korrelation zwischen Teststruktur und Speicherzelle fällt für λ -Werte größer 1.7 zu gering aus, sodass der Ladungseinbau der Speicherzelle unterschätzt wird. λ -Werte kleiner 1.7 führen hinge-

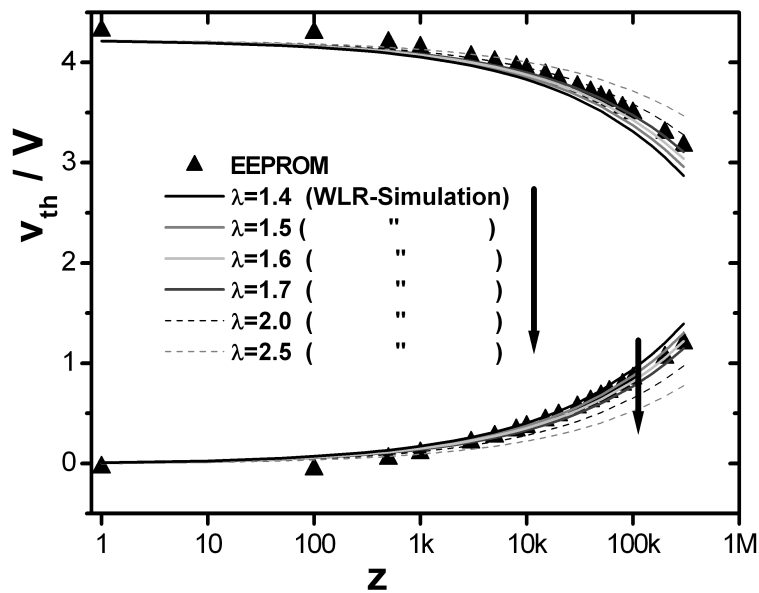


Bild 4.33: Auf der Grundlage von Stressmessungen an Injektor-Teststrukturen (CVS, $\pm 12V$) wurden die Schwellsprungsverläufe in Abhängigkeit von der Zykelzahl simuliert. Der Vergleich mit den Messwerten des 384Bit-Arrays zeigt, dass für λ -Parameter zwischen 1.4 und 1.7 eine gute Übereinstimmung zwischen Simulation und Messung vorliegt. Für λ -Parameter oberhalb von 1.7 wird die Oxiddegradation deutlich unterschätzt.

gen zu einer Überschätzung der Oxiddegradation, woraus zu große Änderungen der Schwellsprungswerte resultieren.

In Bild 4.34 wurde die Änderung der Schwellsprungen für $\lambda = 1.5$ simuliert und mit statistischen Größen verglichen. Hierbei zeigt sich, dass die simulierten Schwellsprungsverläufe vollständig im Korridor zwischen den Minima und Maxima des 384Bit-Arrays verlaufen. Die Simulation liegt für Zykelzahlen unterhalb von 150.000 P/E-Zyklen innerhalb der berechneten Standardabweichung. Oberhalb von 150.000 P/E-Zyklen überschätzt die Simulation die Oxiddegradation. Die Ursache hierfür ist die lineare Näherung, die für die Berechnung der injizierten Ladung verwendet wird. Bei dieser Näherung wurde ein linearer Zusammenhang zwischen injizierter Ladung und Zykelzahl angenommen. Diese Annahme ist näherungsweise für moderate Oxiddegradationen korrekt. Für hohe Oxiddegradationen liegt hingegen kein linearer, sondern ein

4 Untersuchungen zur Oxiddegradation

sublinearer Zusammenhang zwischen beiden Größen vor [135]. Eine Folge hiervon ist die Überschätzung der injizierten Ladung bzw. des Ladungseinbaues, was oberhalb von 200.000 P/E -Zyklen deutlich wird.

In Bild 4.35 wurden die Schwellspannungsverläufe bis 2 Millionen P/E -Zyklen simuliert und zusammen mit Messwerten dargestellt. Eine gute Übereinstimmung zwischen Simulation und Messung kann im programmierten Zustand bis 150.000 P/E -Zyklen und im gelöschten Zustand bis 300.000 P/E -Zyklen festgestellt werden. Damit liegt in dem für die Automobilelektronik relevanten Bereich ($z \leq 150.000$) eine sehr gute Übereinstimmung zwischen Simulation und Messung vor. Die Abweichungen zwischen den Simulationen und Messergebnissen liegen unterhalb von 5 %. Eine deutliche Überschätzung der Oxiddegradation zeigt sich oberhalb von 300.000 P/E -Zyklen. Das Vorhandensein eines Schnittpunktes der simulierten Schwellspannungsverläufe ist ein weiterer Beleg für die Überschätzung des Ladungseinbaues durch das verwendete lineare Modell für Q_{inj} . Die Abweichung zwischen Messungen und Simulation zeigt sich auch in der Bestimmung maximal ausführbarer Zykelzahlen. Wird für die Schwellspannungen eine Detektionsgrenze von 1 V angenommen, um die sich die Schwellspannungen beider Zustände unterscheiden dürfen ($v_{thW} \leq 1$ V), so wird über das WLR-Verfahren eine maximal ausführbare Zykelzahl von $N_{max} \approx 600.000$ gefunden. Die an den Speicherzellen verifizierten Zykelzahlen liegen deutlich darüber, was in Bild 4.35 am Schwellspannungsfenster (rechte Achse) ablesbar ist.

Die Untersuchungen in (Kapitel 4.5.3, Bild 4.28) zeigten eine geringe Spannungsabhängigkeit des Ladungseinbaues der Speicherzelle, sodass mit diesem Simulationsverfahren auch die Spannungsabhängigkeit der Datenwechselstabilität simuliert werden kann. In Bild 4.36 wurde für verschiedene P/E -Spannungen v_{thH} und v_{thL} simuliert und zusammen mit Messwerten als Funktion der Zykelzahl dargestellt. Es zeigt sich, dass der Schwellspannungsverlauf im gelöschten Zustand besser wiedergegeben wird als im programmierten Zustand. Der für die Automobilelektronik interessante Bereich wird auch für verschiedene P/E -Spannungen mit hinreichender Genauigkeit beschrieben. Die Abweichungen zwischen Messung und WLR-Verfahren liegen unter 10 %.

4.6.6 Modelloptimierung

Die Simulationen der Datenwechselstabilität haben gezeigt, dass oberhalb von 150.000 P/E -Zyklen merkliche Abweichungen zwischen Messung und Simulation auftreten. Verursacht wird die Abweichung durch die Linearisierung der injizierten Ladung, die im Bereich großer Zykelzahlen den Ladungsfluss überschätzt.

Der Zusammenhang zwischen injizierter Ladung und Zykelzahl wird im Bereich hoher Zykelzahlen über einen sublinearen Zusammenhang zwischen Q_{inj} und Zykelzahl beschrieben, Gl. 4.24 a) [133]. Diese Gesetzmäßigkeit soll unter Berücksichtigung der Programmier- und Löschspannung verallgemeinert werden Gl. 4.24 b) und im Rahmen des WLR-Verfahrens zum Einsatz kommen.

$$a) \quad Q_{inj}(z) = \alpha z^{\beta} \quad (\beta < 1) \quad b) \quad Q_{inj}(z, V_{P/E}) = \alpha(V_{P/E}) \cdot z^{\beta(V_{P/E})} \quad (4.24)$$

Für die Modellierung der injizierten Ladung wurde diese für verschiedene Programmier-

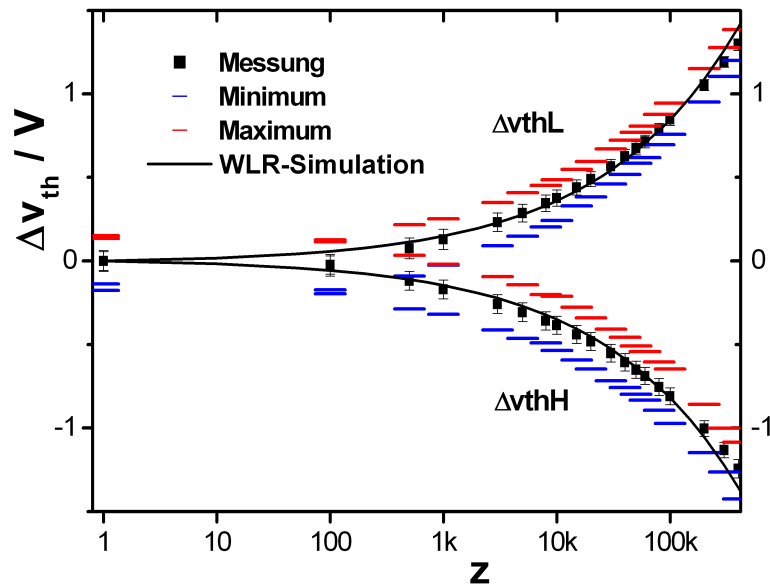


Bild 4.34: Die Änderung der Schwellspannungen wurde für $\lambda=1.5$ simuliert und mit Messwerten verglichen. Neben den gemessenen Mittelwerten und Standardabweichungen des 384Bit-Arrays sind auch die Minima und Maxima dargestellt.

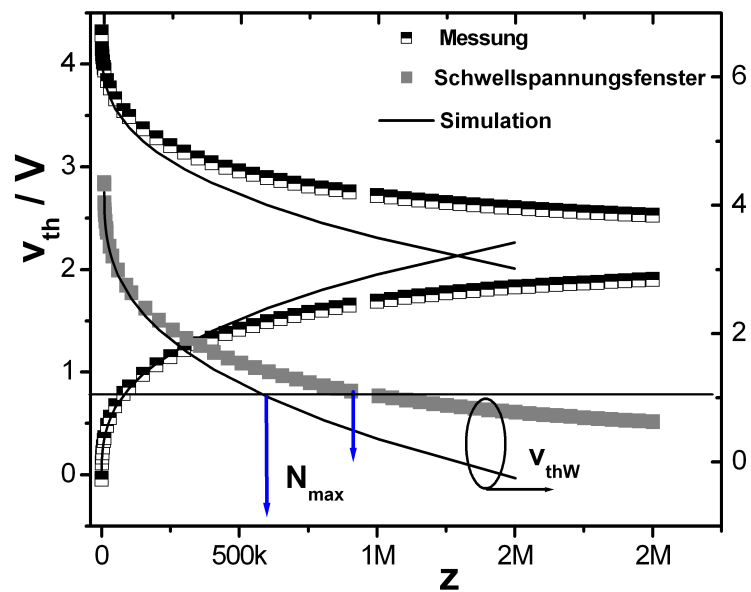


Bild 4.35: Vergleich zwischen WLR-Verfahren und EEPROM-Messergebnissen für hohe Zyklenzahlen. Unterhalb von 200.000 liegt eine gute bis sehr gute Übereinstimmung zwischen Messung und Simulation vor. Für hohe Zyklenzahlen weichen die simulierten Schwellspannungsverläufe deutlich von den Messwerten ab.

4 Untersuchungen zur Oxiddegradation

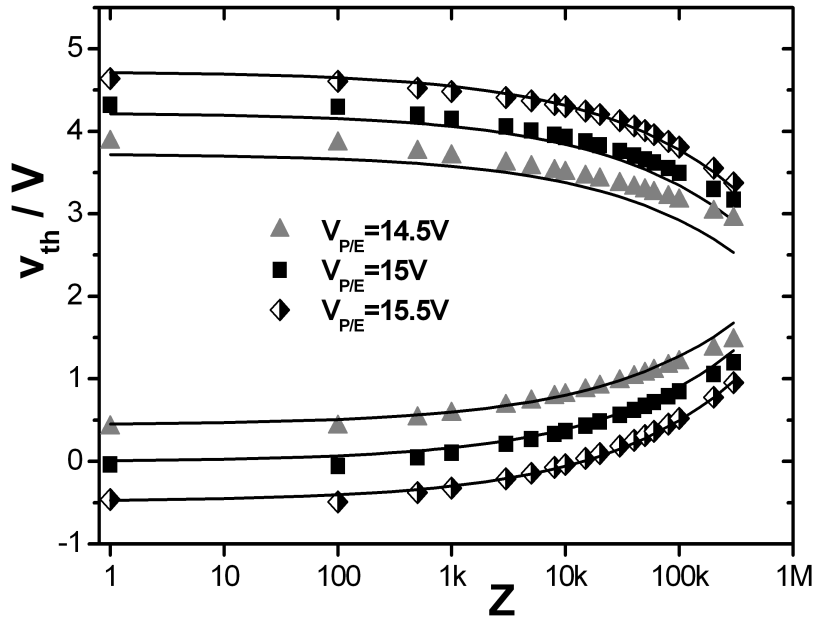


Bild 4.36: Aufgrund der schwachen Spannungsabhängigkeit des Ladungseinbaues können die Schwellspannungsverläufe auch spannungsabhängig simuliert werden. Die Schwellspannungsverläufe wurden für verschiedene P/E-Spannungen simuliert und gemessen. Bis 100.000 P/E-Zyklen weichen die Ergebnisse des WLR-Verfahrens weniger als 10 % von den Messwerten ab.

und Löschspannungen gemessen und in Abhängigkeit von der Zykelzahl dargestellt und mit Gl. 4.24 a) modelliert, Bild 4.37. Es konnte nachgewiesen werden, dass β quasi spannungsunabhängig ist und die Spannungsabhängigkeit der injizierten Ladung auf den Fitparameter α zurückzuführen ist. Die Spannungsabhängigkeit des α -Parameters ist durch einen linearen Zusammenhang beschreibbar, der aus Bild 4.38 hervorgeht. Die injizierte Ladung wurde nach 100.000 P/E-Zyklen bestimmt und über der P/E-Spannung aufgetragen.

$$Q_{inj}(V_{P/E}) \propto V_{P/E} \Rightarrow Q_{inj}(V_{P/E}) = a^* + b^*(V_{P/E}), \quad (4.25)$$

mit a^* , b^* als Fitparameter. Wird die Potenzgesetzmäßigkeit Gl. 4.24 a) mit Gl. 4.25 kombiniert und ferner berücksichtigt, dass die injizierte Ladung proportional zu $(2C_{CG}/A_{inj})$ ist, ergibt sich folgender Ausdruck:

$$Q_{inj} = \frac{2C_{CG}}{A_{inj}} (a + b \cdot V_{P/E}) z^c \quad (4.26)$$

mit a , b und c als spannungsunabhängigen Fitparametern.

In Bild 4.39 ist die injizierte Ladung in Abhängigkeit von z und $V_{P/E}$ dargestellt und mit der Gesetzmäßigkeit Gl. 4.26 modelliert. Das Modell beschreibt das sublineare Verhalten zwischen injizierter Ladung und Zykelzahl sowie die Abhängigkeit von der Programmier- und Löschspannung. Im Gegensatz zum linearen Modell, das bereits ab

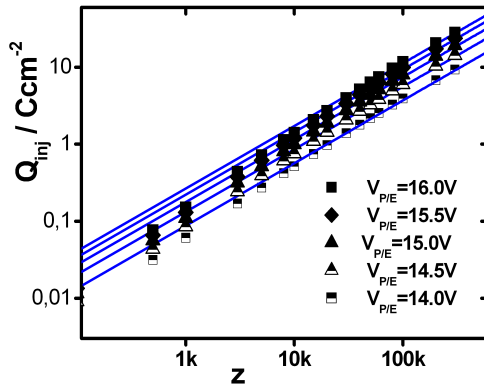


Bild 4.37: Darstellung der injizierten Ladung $Q_{inj}(z)$ für verschiedene Programmier- und Löschspannungen.

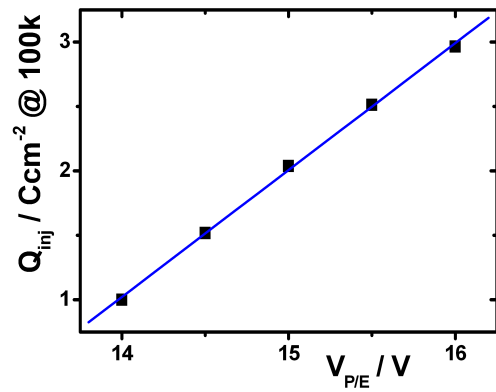


Bild 4.38: Die injizierte Ladung wurde nach 100.000 P/E-Zyklen bestimmt und über der P/E-Spannung aufgetragen.

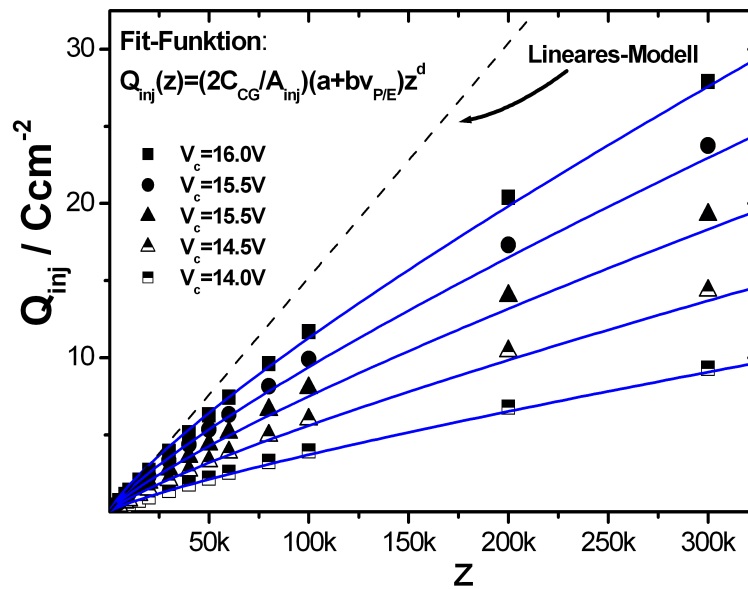


Bild 4.39: Zeigt die injizierte Ladung in Abhängigkeit von der Zykelzahl für verschiedene P/E-Spannungen. Die Gleichung 4.26 ermöglicht eine Modellierung der injizierten Ladung bis 300.000 P/E-Zyklen.

4 Untersuchungen zur Oxiddegradation

50.000 P/E-Zyklen deutlich von den Messwerten abweicht, gelingt über das sublineare Modell eine Modellierung bis 300.000 P/E-Zyklen.

Die Anwendung des sublinearen Modells im Rahmen des WLR-Verfahrens ist in Bild 4.40 zu sehen. Zusätzlich sind EEPROM-Messergebnisse und die Simulationsergebnisse des linearen Modells dargestellt. Die lineare Näherung führt ab 100.000 P/E-Zyklen zur Überschätzung der Oxidalterung, was auf zu niedrige Schwellspannungswerte hinausläuft. Im Vergleich hierzu liefert das „sublineare“ Modell eine gute Übereinstimmung zwischen Simulation und EEPROM-Werten bis 300.000 P/E-Zyklen mit einem Fehler, der deutlich unter 5 % liegt.

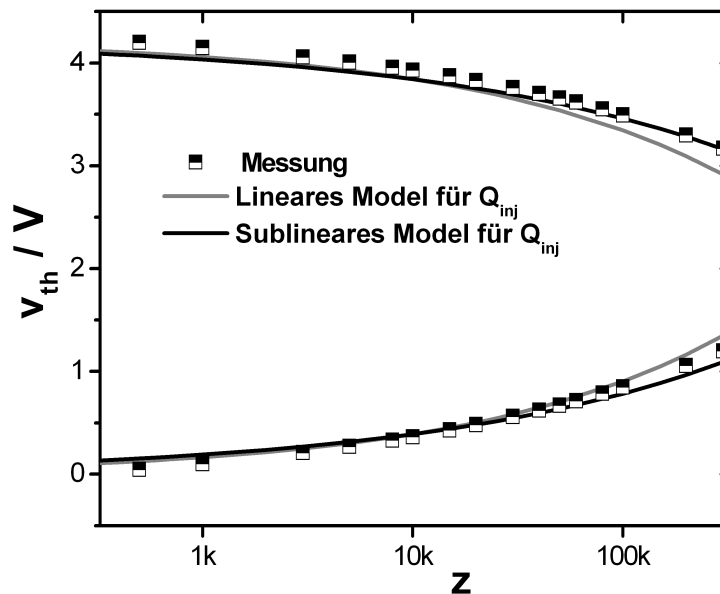


Bild 4.40: Simulation der Datenwechselstabilität mithilfe verschiedener Modelle für $Q_{inj}(z)$. Die Schwellspannungen wurden in Abhängigkeit von der Zykelzahl dargestellt und mit den Messwerten des EEPROM-Arrays verglichen.

5 Zusammenfassung und Diskussion

Die Modellierung der Speicherzelle in Kapitel 2 und 3 hat gezeigt, dass das Programmier- und Löschverhalten mithilfe des vorgestellten Kapazitätsmodells und der Fowler-Nordheim-Gleichung qualitativ richtig wiedergegeben wird. Durch die Einbeziehung von Prozess- und Layoutparametern konnte das Modell über die Fowler-Nordheim-Größen A und B an die Messdaten angepasst werden, um eine quantitative Beschreibung der Speicherzellen zu ermöglichen, die in guter Übereinstimmung mit den Messungen steht.

Eine weitere Zielsetzung war die Untersuchung des Ladungseinbaues bzw. der Oxiddegradation unter verschiedenen Stresskonditionen. Es wurden zwei Messverfahren implementiert, mit denen die Speicherzellen und die Injektorteststrukturen untersucht werden konnten. Die Oxidladung wurde in Abhängigkeit von der injizierten Ladung gemessen und ließ sich mit einem Potenzgesetz von Papadas et al. modellieren. Die hierüber gefundenen Modellparameter n und K ermöglichen die Beurteilung der Oxiddegradation und ferner die Auffindung einer Korrelation zwischen beiden Bauteilen. Diese Korrelation ist auf die geringe Abhängigkeit des n -Parameters vom verwendeten Stressmodus zurückzuführen und setzt den Ladungseinbau der Speicherzelle in Relation zum Ladungseinbau der verwendeten Teststrukturen. Dieser Zusammenhang stellt die wichtigste Erkenntnis der vorliegenden Arbeit dar.

Die experimentellen und theoretischen Untersuchungen zum Ladungseinbau waren Ausgangspunkt für die Entwicklung eines WLR-Tests zur Bestimmung der Zykelfestigkeit und Oxidqualität. Die zentrale Messgröße des Tests war die Oxidladung, die in Abhängigkeit von der injizierten Ladung gemessen wird: $Q_{ox}(Q_{inj})$. Der funktionale Zusammenhang zwischen Oxidladung und Ladungsdurchfluss wurde in den Mittelpunkt der Untersuchungen gestellt, da die Zykelfestigkeit in direkter Korrelation mit den Oxidladungen bzw. dem Oxidladungseinbau steht. Auf der Basis von Stressmessungen können mit dem in Kapitel 4 entwickelten Verfahren die Schwellspannungsverläufe als Funktion der Zykelzahl simuliert werden. In dem für Automobilelektronik wichtigen Bereich ($z \leq 150.000$) wurde eine sehr gute Übereinstimmung zwischen Messung und Simulation erzielt. Erst oberhalb von 150.000 P/E -Zyklen wurden Abweichungen festgestellt. Um dieses Problem zu lösen, wurde für die injizierte Ladung ein Modell mit sublinearem Verhalten entwickelt, mit dem die Genauigkeit der Simulationsergebnisse gesteigert werden konnte. Wie in Bild 5.1 schematisch gezeigt wird, erfordert die Bestimmung des Ladungseinbaues nur wenige Messpunkte, da weitere Messpunkte extrapoliert werden können. Um das Beschleunigungsvermögen des WLR-Testes zu berechnen, kann ein Beschleunigungsfaktor b_{WLR} errechnet werden, der die Testzeit des WLR-Tests mit der Testzeit am EEPROM-Array in Relation setzt. Wird für die Zeitdauer des WLR-Testes eine Zeit von ca. 20 s angenommen, kann aus der Zeitdauer für eine

5 Zusammenfassung und Diskussion

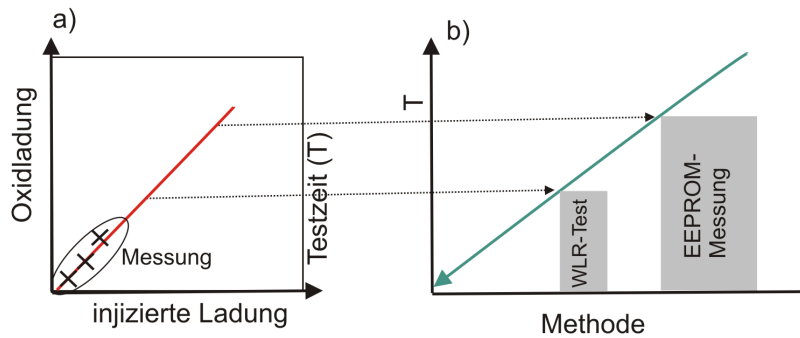


Bild 5.1: a) Wenige Messpunkte sind ausreichend, um den Oxidladungseinbau der Teststrukturen zu extrapolieren. b) Die Datenwechselstabilität kann mit den extrahierten Oxidladungen und den Modellen aus Kapitel 3 und 4 simuliert und bewertet werden.

spezifizierte Anzahl von P/E -Zyklen der Beschleunigungsfaktor berechnet werden:

$$b_{WLR} = \frac{2z \cdot t_p}{t_T}, \quad (5.1)$$

wobei t_p die Pulslänge und z die Zykelzahl bezeichnet. Für die Ausführung von 100.000 P/E -Zyklen werden etwa 40 Minuten benötigt. Hiermit ergibt sich ein Beschleunigungsfaktor ≥ 120 . Durch diesen deutlichen Zeitgewinn gegenüber der direkten Messung an Speicherzellen, Teilbild b), kann der Stichprobenumfang pro Wafer und Charge signifikant erweitert werden. In Hinblick auf das Optimierungspotenzial des WLR-Tests stellt sich die Frage, ob der Test weiter beschleunigt und die verwendete Modellierung optimiert werden kann. In diesem Kontext sind Messungen bei erhöhten Temperaturen und Spannungen von Interesse. Eine weitere Beschleunigung der Oxiddegradation durch die Erhöhung der Stressspannung ≥ 12 V ist begrenzt möglich. Der Trapping-yield, also das Verhältnis aus Oxidladung und injizierter Ladung ($\Delta Q_{ox}/Q_{inj}$), ist spannungsabhängig und nimmt mit zunehmender Stressspannung deutlich zu [134]. Allerdings haben Untersuchungen im Rahmen dieser Arbeit gezeigt, dass der Trapping-yield mit steigender Spannung ein Sättigungsverhalten aufweist und die Oxiddegradation ab einer Spannung (≥ 12 V) nur noch marginal beschleunigbar ist. Ferner birgt eine weitere Erhöhung der Spannung die Gefahr in sich, dass sich der Degradationsmechanismus und damit verbunden der Ausfallmechanismus ändert.

Eine weitere Option zur Beschleunigung der Oxidalterung ist die Erhöhung der Umgebungstemperatur, was deutlich aus den Untersuchungen in Kapitel 4 hervorging. Daher sollte grundsätzlich eine weitere Beschleunigung des WLR-Testes durch Temperaturerhöhung möglich sein. In zukünftigen Untersuchungen müsste daher die Gültigkeit der Korrelation zwischen Speicherzelle und Teststruktur verifiziert bzw. modifizieren werden. Weiterhin wäre eine Verbesserung des Modells zur Simulation der Schwellspannungen erforderlich, da der Auslesetransistor zahlreiche Temperaturabhängigkeiten aufweist. Eine Modellentwicklung basierend auf einem Kompaktmodell (BSIM3) und die Implementierung in einen geeigneten Device-Simulator wären eine Möglichkeit, diese Modifikationen zu realisieren [40,133,135].

Literaturverzeichnis

- [1] Internetseite: Bundesministerium für Bildung und Forschung „Elektronik und Elektroniksysteme „made in Germany “, *www.bmbf.de*, 20.05.2010
- [2] Dr.-Ing. K. -J. Amsler, Dr.-Ing. J. Fetzler, Dr.-Ing D. Lederer, Dr. M. Erben “Sicherheitsgerechte Entwicklungsprozesse- alles neu geregelt “ *Vector Consulting GmbH, Stuttgart, Quelle: Internet*
- [3] R. Ingenbleek und T. Riepl “Integration of Micromechanic Sensors, Actuators and Miniaturized High Temperature Electronics in Advanced Transmission Systems“ *in Proceeding of Micro. Tec. 2000, Hannover, Sep. 2000*
- [4] R. Groppo, W. Wondrak, G. Lugert und T. Riepl , “Moving towards high temperature electronics for automotive applications“ *HITEN 2001, Book of Abstracts, Oslo, 5-6 Juni, pp. 105-111, 2001*
- [5] D. Kahng and S. M. Sze, “A floating gate and its application to memory devices“ *Bell Syst. Tech. J.*, vol. 46, p. 1288, 1967
- [6] W. D. Brown, J. E. Brewer “Nonvolatile Semiconductor Memory Technology, A comprehensive guide to understanding and using NVSM devices “ *IEEE Press, 1998*
- [7] H. A. R. Wegner, A. J. Lincol, H. C. Pao, M. R. O’Connell, H. Lawrence, R. E. Oleksiak “The variable threshold transistor, a new electrically alterable, non destructive read-only storage device“ *IEEE IEDM Tech. Dig., Washington D. C., 1967*
- [8] D. Frohman-Bentchkowsky “A fully decoded 2048-bit electrically programmable MOS-ROM “ *IEEE ISSCC Dig. Tech. Pap.*, pp. 80-81, 1971
- [9] D. Frohman-Bentchkowsky “Memory behaviour in a floating gate avalanche injection MOS (FAMOS) structure “ *Appl. Phys. Lett.*, vol. 18, pp. 332-334, 1971
- [10] W. Johnson, G. Peregos, A. Renninger, G. Kuhn and T. Ranganath “A 16Kb electrically erasable nonvolatile memory“ *IEEE ISSCC Dig. Tech. Pap.*, p. 152, 1980
- [11] J. Yeargain and K. Kuo “A high density floating gate EEPROM cell“ *IEEE IEDM Tech. Dig.*, p. 24, 1981
- [12] D. Guterman, I.H. Rimawi, T. Chiu, R.D. Halvorson and D.J. McElroy “An electrically alterable nonvolatile memory cell using a floating gate structure “ *IEEE Trans. Elect. Dev.*, vol. ED-26, pp. 576-586, 1979
- [13] B. Rössel and R. Müller “Electrically erasable and reprogrammable read-only memory using the n-channel SIMOS one-transistor cell“ *IEEE Trans. Elect. Dev.*, vol. ED-24, p. 806, 1979

- [14] R. Benz, E. Camerlenghi, A. Modelli and A. Visconti "Introduction to Flash Memory " *Proceedings of the IEEE*, vol. 91, no. 4, August 2003
- [15] E. Harari, L. Schmitz, B. Troutman and S. Wang "A 256bit non-volatile static RAM" *IEEE ISSCC Dig. Tech. Pap.*, pp. 108-109, 1978
- [16] F. Masuoaka, M. Assano, H. Iwahashi, T. Komuro and S. Tanaka "A new Flash EEPROM cell using triple polysilicon technology " *IEEE IEDM Tech. Dig.* pp. 464-467, 1984
- [17] F. Masuoaka et al. "A 256K flash EEPROM using triple polysilicon technology" *IEEE ISSCC Dig. Tech. Pap.*, pp. 168-169, 1985
- [18] F. Masuoaka, M. Momodoni, Y. Iwata and R. Shirola "New ultra high density EPROM and Flash EEPROM cell with NAND structure cell" *IEEE IEDM Tech. Dig.*, pp. 552-555, 1987
- [19] E. Takeda, H. Kume, T. Toyabe and S. Asai "Submicrometer MOSFET structure for minimizing hot carrier generation" *IEEE Transaction Elect. Dev.*, vol. ED-29, pp. 611-618, 1982
- [20] C. Hu "Lucky electron of hot electron emission " *IEEE IEDM Tech. Dig.*, p.22, 1979
- [21] P.K. Ko, R.S Müller and C. Hu "A unified model for hot electron currents in MOSFET's" *IEEE IEDM Tech. Dig.*, pp. 600-603, 1981
- [22] G. S. May und S. M. Sze " Fundamentals of Semiconductor Fabrication " *Wiley International* 2004
- [23] Reisch, "Halbleiterbauelemente" *Springerverlag* 2005
- [24] T. Giebel "Grundlagen der CMOS-Technologie" *Teubner Verlag* 2002
- [25] J. Peper "Zeitoptimierte Verfahren zur Bestimmung des Datenerhalts und der Wiederbeschreibbarkeit von Doppel-Poly EEPROM Zellen - Korrelation zu technologischen und designabhängigen Parametern." *Dissertation Fachbereich Elektrotechnik der Gerhard- Mercator- Universität Gesamthochschule Duisburg*, 2001
- [26] C. Plossu, J. M. Voisin, B. Bos, C. Raynaud, R. Bouchakour, P. Boivin, B. Balland "Dynamic stress of thin tunnel oxides: a way to emulate a single EEPROM cell programming function", *Journal of Non-Crystalline Solids*, vol. 245, pp. 85-91, 1999
- [27] P. Pavan, R. Benz, P. Olivo and E. Zanoni " Flash Memory Cell- An Overview " *Proceedings of the IEEE*, vol. 85, no. 8, August 1997
- [28] Ching-Yuang Wu, Chiou-Feng Cheng "Physical model for characterization and simulating a FLOTOX EEPROM device" *Solids-State Electronics*, vol. 35, no. 5, pp. 705-716, 1992
- [29] S. Sze, "Physics of Semiconductor Devices" *Wiley-Interscience*, 1981
- [30] D. Ielmini, A. S. Spinelli, A. L. Lacaita, A. Modelli „Modeling of anomalous SILC in flash memories based on tunneling at multiple defects" *Solid State Electronics*, vol. 46, pp. 1749-1756, 2002

- [31] D. Ielmini et al „A statistical model for SILC in Flash memories“ *IEEE Transaction on Electron devices*, vol. 49, no.11, pp. 1955-1961, 2002
- [32] F. Schuler, R. Degraeve, P. Hendickx, D. Wellekens „Physical description of anomalous charge loss in floating gate based NVM's and identification of its dominant parameter“ *IEEE International Reliability Physics Symposium, Dallas Texas 2002*
- [33] M. Herman, A. Schenk „Field and high-temperature dependence of the long term charge loss in erasable programmable read only memories: Measurements and modeling“ *J. Appl. Phys.*, vol. 77, no. 9, 1995
- [34] L. Larcher, S. Bertulu and P. Pavan „SILC effect on EERPOM memory cell reliability“ *In Proc. ESSDERC, Edingburgh UK*, p. 495, 1994
- [35] R. H. Fowler and L. Nordheim “Electron emission in intense electric films “ *Royal Soc. London*, vol. 119, pp. 173-181, 1928
- [36] Siegfried Flügge “ Rechenmethoden der Quantenmechanik “, *Springerverlag*, 6 Auflage, 1999
- [37] M. Lenzlinger and E. H. Snow “Fowler-Nordheim Tunneling into Thermally Grown SiO₂ “ *J. Appl. Phys.* vol. 40, no. 1, 1969
- [38] G. Pananakakis, G. Ghibaudo and R. Kies “Temperature dependence of the Fowler-Nordheim current in metal-oxide-degenerate semiconductor structures“ *J. Appl. Phys*, vol. 78, no. 4, pp. 2635-2641, 1995
- [39] Y. Zahi, R. Laffont, F. Lalande, R. Bouchakour “Tunneling injection temperature dependence in EEPROM cell“ *Journal of Non-Crystalline Solids*, vol. 353, pp. 648-652, 2007
- [40] F. Schuler, O. Kowarik, K. Hoffmann “ BSIM3v3 based Degradation Compact Model for Circuit Simulation of Non-Volatile Flash Memories “ *ICSE Proceedings*, 1998
- [41] J. A. Lopez-Villanueva, J. A. Jimenez-Tejada, P. Cartujo, J. Bausells and J. E. Carceller “Analysis of the effect of constant-current Fowler-Nordheim-tunneling injection with charge trapping inside the potential barrier“ *J. Appl. Phys.*, vol. 70, no. 7, 1991
- [42] C. Papadas, G. Ghibaudo, G. Pananakakis and C. Riva “Numerical transient simulation of the programming window degradation in FLOTOX EEPROM cell“ *Solid-State Electronics* , vol. 36, no. 9, pp. 1303-1311, 1993
- [43] J. S. Witters, G. Groeseneken, H. E. Maes “Degradation of tunnel-oxide Floating-Gate EEPROM Devices and the Correlation with high field current induced degradation of thin gate oxides “ *IEEE Trans. Electron Devices*, vol. 36, no. 9, pp. 1663-1682, 1989
- [44] D. M. Fleetwood, “ Border Traps in MOS Devices “*IEEE Trans. Nucl. Sci.*, vol. 39, no. 2, pp. 269-271, April 1992
- [45] “A. Spinelli, A. L. Lacaita, M. A. Rigamonti and G. Ghidini „Experimental method for the determination of the energy distribution of stress-induced oxide traps“ *IEEE Electron Device Lett.*, vol. 20, no. 3, pp. 106-108, 1999

- [46] H. Flietner, "Passivity and electronic properties of the silicon/silicondioxide interface" *Material Science Forum*, vol. 185-188, pp. 73-82, 1995
- [47] W. Flüssel, M. Schmidt, H. Angermann, G. Mende, H. Flietner "Defects at the Si/SiO₂ interface: their nature and behaviour in technological process and stress" *Nuclear Instr. and methods in Phys. Research A*, vol. 377, issue 2-3, pp. 177-183, 1996
- [48] J. K. Rundra, W. B. Fowler and F. J. Feigl "Model for the E₂'-center in alpha quartz" *Phys. Rev. Lett.*, vol. 55, no. 23, pp. 2614-2617, December 1985
- [49] F. Beug "Charakterisierung von EEPROM Tunneloxiden mittels transients Strom- und Kapazitätsmessungen" *Dissertation, Universität Hannover*, 2005
- [50] L. D. Thanh and P. Balk "Elimination and generation of Si-SiO₂ Interface traps by low temperature hydrogen annealing" *J. Electroch. Soc.*, vol. 135, no. 7, pp. 1797-1801, 1988
- [51] J. H. Stathis "Electrically detected magnetic resonance study of stress-induced leakage current in thin SiO₂" *Appl. Phys. Lett.* vol. 68, no. 12, pp. 1669-1671, 1996
- [52] D. J. DiMaria, E. Cartier and D. Arnold "Impact Ionization, Trap generation, degradation and breakdown in silicondioxide films on silicon" *J. Appl. Phys.*, vol. 73, no. 7, pp. 3367-3384, 1993
- [53] I. Yoshii, K. Hama and K. Hashimoto "Role of hydrogen at poly Si/SiO₂ interface in trap generation by substrate hot electron" *Proceedings of IRPS*, pp. 136-140, 1992
- [54] K. F. Schuegraf and C. Hu "Metal-Oxide-semiconductor field-effect transistors substrate current during fowler-nordheim tunneling stress and silicon-dioxide reliability" *J. Appl. Phys.*, vol. 76, no. 6, pp. 3695-3700, 1994
- [55] K. F. Schuegraf and C. Hu "Hole injection SiO₂ breakdown model for very low lifetime extrapolation" *IEEE Trans. Electron Devices*, vol. 41, no. 5, pp. 761-767, 1994
- [56] D.J. DiMaria and J.W. Stasiak "Trap generation in silicon dioxide produced by hot electrons" *IEEE J. Appl. Phys.*, vol. 56, no. 6, pp. 2342-2356, 1989
- [57] R. Degraeve et al. "Analytical Percolation Model for Predicting Anomalous Charge Loss in Flash Memories" *IEEE Trans. Electron Devices*, vol. 51, no. 9, 2004
- [58] J. H. Stathis "Percolation models for gate oxide breakdown" *J. Appl. Phys.*, vol. 86, no.10, pp.3367-3384,1999
- [59] B. E. Deal "Standardized terminology for oxide charges associated with thermally oxidized silicon" *IEEE Trans. Electron Devices*, vol. ED-27, no. 3, pp. 606-608, 1980
- [60] D. J. DiMaria, *Proc. International Topical Conf. on Physics of SiO₂ and its interfaces*, Yorktown Heights, New York, 1978
- [61] N. H. Ning and H. N. Yu "Optically induced injection of hot electrons into SiO₂" *J. Appl. Phys.*, vol. 45, no. 12, pp. 5373-5378 1974

- [62] M. S. Liang and C. Hu “Electron trapping in very thin thermal silicon dioxides“ *IEDM, 1981*
- [63] C. Busseret, N. Baboux, C. Plossu, S. Burignat, P. Boivin “Quantitative study of charge trapping in SiO_2 during bipolar Fowler-Nordheim injection “ *Journal of Non-Cryst. Solids, vol. 322, pp. 191-198, 2003*
- [64] Nolting “Grundkurs Theoretische Physik, Band 5 “ *Springer Verlag 2001*
- [65] M. Houssa, A. Stesmans, M. Naili, M. M. Heyns “Charge trapping in very thin high-permittivity gate dielectric layers “ *Appl. Phys. Letters, vol. 77, no. 9, 2000*
- [66] C. Papadas, G. Ghibaudo, G. Pananakakis “Oxide reliability criterion for the evaluation of the endurance performance of electrically erasable programmable read only memories “ *J. Appl. Phys., vol. 71, no. 9, 1 May 1992*
- [67] A. Scarpa, G. Pananakakis, G. Ghibaudo, A. Paccagnella, “On the degradation kinetics of thin oxide layers “ *Solid-State Electronics 43, pp. 221-227, 1999*
- [68] G. Pananakakis, G. Ghibaudo, G. Papadas, C. Vincent, E. Kies “Generalized trapping kinetic model for the oxide degradation after Fowler-Nordheim uniform gate Stress“ *J. Appl. Phys. 82(5), 1997*
- [69] D. Gogl, B. Burbach, H. L. Fiedler, M. Verbeck and C. Zimmermann “Single Poly EEPROM Cell in SIMOX Technology for High-Temperature Applications up to 250°C “ *IEEE Elec. Dev. Letters., vol. 18, no. 11, 1997*
- [70] R. Benz, E. Camerlenghi, A. Modelli and A. Visconti “Introduction to Flash Memory “ *Proceedings of the IEEE, vol. 91, no. 4, August 2003*
- [71] R. Benz, L. Baldi, E. Camerlenghi and P. Cappelletti “An overview of flash memories (STMicroelectronics)“ “*International workshop on emerging non volatile memories, 14 Sept. 2007* “
- [72] P. Cappelletti, C. Golla, P. Olivo and E. Zanoni „Flash Memories“ *Kluwer Academic Publishers, Massachusetts, 1999*
- [73] C. Papadas, G. Pananakakis, G. Ghibaudo, C. Riva, F. Pio and P. Ghezzi “Modeling of the intrinsic retention characteristics of FLOTOX EEPROM Cells under elevated temperature conditions“ *IEEE Transaction on Electron Devices, vol. 42, no. 4, 1995*
- [74] E. Vincent, C. Papadas, C. Riva, F. Pio, and G. Ghibaudo „On the charge build-up mechanism in very thin insulator layers“ *In Proc. of ESSDERC, Edinburgh UK, pp. 495-498, 1994*
- [75] K. F. Schuegraf and C. Hu „Defect breakdown lifetime prediction of thin SiO_2 at low voltages “ *In Proc. of IEEE Semiconductor Interface Special. Conf., Ford Lauderdale, p. 93, 1993*
- [76] Y. -B. Park and D. K. Schroeder “Degradation of thin tunnel gate oxide under constant fowler-nordheim current stress for a flash EEPROM“ *IEEE Transaction on Electron Devices, vol. 45, no. 6, pp. 1361-1368 1998*
- [77] L. Bergmann, C. Schäfer, T. Dortfmüller, W. T. Hering, K. Stierstadt „Lehrbuch der Experimentalphysik, Bd. 1 Mechanik, Relativität, Wärme “ *De Gruyter Verlag, 11 Auflage*

- [78] J. V. Manca et al. „The arrhenius relation for electronics in extreme temperature conditions “ *Third conference on high temperature electronics*, pp. 29-32, 1999
- [79] B. De Salvo, G. Ghibaudo, G. Pananakakis, B. Guillaumot “Investigation of low field and high temperature SiO_2 and ONO leakage current using the floating gate technique “ *Journal of Non-Crystalline Solids*, vol 245, pp. 104-109, 1999
- [80] B. De Salvo et al. “Experimental and Theoretical Investigation of Nonvolatile Memory Data-Retention“ *IEEE Trans. Electron Devices*, vol. 46, no. 7, 1999
- [81] P. Canet, R. Bouchakour, N. Harabech, Ph. Boivin, J. M. Mirabel “EEPROM programming study -time and degradation aspects“ *Circuits and systems, ISCAS 2001*
- [82] P. Canet, R. Bouchakour, N. Harabech, Ph. Boivin, J. M. Mirabel, C. Plossu “Study of signal programming to improve EEPROM cell reliability“ *IEEE Midwest Symp. on Circuits and Systems*, vol. 3, pp. 1144-1147, 2000
- [83] V. Bouquet, P. Canet, F. Lalande, J. Devin, B. Leconte, N. Mariema “ Variation of flash memory threshold voltage correlated with applied voltage slope in Fowler Nordheim erase mode“ *Proceedings of Ph.D. Research In Micro-Electronics and Electronics, Lausanne, Switzerland, July 25-28 2005*, vol. 1, pp. 86-88, 2005
- [84] Z. A. Weinberg *J. Appl. Phys.*, vol. 53, 5052, 1982
- [85] Z. A. Weinberg *Solid-State Electron.*, vol. 22, 1977
- [86] *Pers. Mitteilung von Dipl. Phys. Jesco Möller, Elmos AG, 2009*
- [87] A. Concannon, S. Keeney, A. Mathewson, R. Benz, and C. Lombardi „ Two-Dimensional Numerical Analysis of Floating-Gate EEPROM Device“ *IEEE Trans. on Electron Devices*, ED-40 (7), pp. 1258-1262, 1993
- [88] C. Papadas, G. Ghibaudo, P. Mortini and G. Pananakakis “ Analytical modelling of the programmed window in FLOTOX EEPROM cells“ *Electronics Letters*, vol. 29, no. 1, 1993
- [89] R. Benz, D. Cantarelli und P. Cappelletti “Transient Analysis of the current in EEPROM cell“ *IEEE Transaction on Electron Devices*, Bd. 37, pp. 1081-1086, 1990
- [90] M. Itsumi “Positive and negative charging of thermally grown SiO_2 induced by fowler-nordheim emission “ *J. Appl. Phys.* vol. 52, no.5, pp. 3491-3497, 1981
- [91] S. Aritome, R. Shirota, G. Hemink, T. Endoh and F. Masuoka “Reliability Issues of Flash Memory Cells“ *Proceedings of the IEEE*, vol. 82, no. 5, 1993
- [92] S. Croci et al., “Effect of some technological parameters on Fowler-Nordheim injection through oxides for non-volatile memories “ *Journal of Non-Crystalline Solids*, vol. 280, issue 1-3, 2001
- [93] B. E. Deal and M. Sklar, “Thermal oxidation of heavily doped silicon“ *J. Electrochem. Soc.*, vol. 112, pp. 430-435, 1965
- [94] S. A. Campbell “The Science and Engineering of Microelectronic Fabrication“ *Oxford University Press*, 2001

- [95] J. Razafindramora, P. Canet, F. Lalande, R. Bouchakour, R. Bouchakour, P. Boivin, J. -M. Mirabel "Deterioration of stressed oxides: Application to the prediction of non-volatile memory cell endurance" *Journal of Non-Crystalline Solids*, vol. 351, pp. 1878-1884, 2005
- [96] J. Postel-Pellerin, F. Lalande, P. Canet, S. Boutahar, P. Pizzuto, A. Regnier "Impact of stress on fowler-nordheim parameters effects on EEPROM threshold voltages" *Journal of Non-Crystalline Solids*, 353, pp. 610-614, 2007
- [97] D. J. DiMaria "Determination of insulator bulk trapped charge densities and centroids from photocurrent-voltage characteristics of MOS structures" *Journal of Appl. Phys.*, vol. 47, no. 9, 1976
- [98] Y. Nissan-Cohen, J. Shappir and D. Frohman-Bentchkowsky "Characterization of simultaneous bulk and interface high-field trapping effects in SiO_2 " *IEDM*, 1983
- [99] C. Papadas, G. Pio, G. Pananakakis, F. Montserie, P. Mortini, C. Riva *Microelectron. Reliab.* vol. 37, pp. 495-505, 1994
- [100] G. Ghibaudo, G. Pananakakis, R. Kies, E. Vincent, C. Papadas "Accelerated dielectric breakdown and wear out standard testing methods and structures for reliability evaluation of thin oxides" *Microelectronics Reliability*, vol.39, no. 2, pp. 597-613, 1999
- [101] R. Kies, T. Egilsson, G. Ghibaudo and G. Pananakakis "A methode for the assessment of oxide charge density and centroid in metal-oxide-semiconductor structures after uniform gate stress" *Appl. Phys. Lett.*, vol. 68, no. 26, 1996
- [102] R. Kies, T. Egilsson, G. Ghibaudo and G. Pananakakis "Assessment of oxide charge density and centroid from fowler-nordheim derivative characteristics in MOS structures after uniform gate stress" *Microelectron. Reliab.*, vol. 36, no. 11/12, pp. 1619-1622, 1996
- [103] S. Brugler, P. Jespers „Charge pumping in MOS devices“ *IEEE Trans. Electron Devices*, vol. 16, pp. 297-302, 1969
- [104] C. Nguyen, G. Ghibaudo, F. Balestra *Phys. Stat. Solids(a)*, 126, 553, 1991
- [105] P. J. McWhorter and P. S Winokur "Simple technique for seperating the effect of interface traps and trapped oxide charge in metal-oxide-semiconductor transistors" *Appl. Phys. Lett.*, vol. 48, no. 2, 1986
- [106] A. Aal "A comparison between V-Ramp TDDB techniques for reliability Evaluation" *IRW FINAL REPORT 2008*
- [107] A. Aal "fWLR supported Process Development based on V- and J-Ramp Stress Tests" *IRW FINAL REPORT 2009*
- [108] R. M. Anderson and D. R. Kerr "Evidence for surface asperity mechanism of conductivity in oxide grown on polycrystalline silicon" *J. Appl. Phys.*, vol. 48, no. 11, pp. 4834-4836, 1977
- [109] D.J. DiMaria and D.R. Kerr „ Interface effects and high conductivity in oxides grown from polycrystalline silicon " *Appl. Phys. Lett.*, vol. 27, no. 9, pp. 505-507, 1975

- [110] E. Vincent, S. Bruyere, C. Papadas, P. Mortini " Dielectric reliability in deep-submicron technologies: From thin to ultrathin oxides" *Microelectron. Reliab.* vol. 37, no. 10/11, pp. 1499-1509, 1997
- [111] S. D. Brorson, D. J. DiMaria, M. V. Fischetti, F. L. Pesavento, P. M. Solomon, D. W. Kwong „Direct measurement on energy distribution of hot electrons in silicon dioxide “ *J. Appl. Phys.* 58, 1302, 1985
- [112] I.-C.Chen, S. Holland, C. Hu, „Oxide breakdown dependence on thickness and hole current-enhanced reliability of ultra thin oxides“ *IEEE Int. Electron Dev. Meeting Tech. Digest*, pp. 660-663, 1986
- [113] E. Vincent C. Papadas, and G. Ghibaudo "Temperature dependence of charge build-up mechanism and breakdown phenomena in thin oxides under fowler-nordheim injection " *Solid-State Electronics*, vol. 41, no. 7, pp. 1001-1004, 1997
- [114] C. Papadas, G. Ghibaudo, G. Pananakakis, C. Riva, P. Ghezzi "Programming window degradation in FLOTOX EEPROM cells degradation" *Microelectronic Engineering*, vol. 15, pp. 621-624, 1991
- [115] A. Bhattacharyya " Modelling of write/erase and charge retention of floating gate EEPROM device" *Solid-State Electron*, 1984, vol. 27, no. 10, pp. 899-906
- [116] P. Canet, R. Bouchakour, J. Razafindramora, F. Lalande, J. M. Mirabel "Very fast EEPROM Erasing study" *ESSCIRC*, 2002
- [117] A. Cester, A. Paccagnella, G. Ghidini "Stress induced leakage current under pulsed voltage stress" *Solid-State Electronics*, vol. 46, no. 3, pp. 399-405, 2002
- [118] E. Rosenbaum, Z. Liu, C. Hu, „ Silicon dioxide breakdown lifetime enhancement under bipolar bias conditions “ *IEEE Trans. Electron Dev.*, vol.40, no. 12, pp. 2287-2295, 1993
- [119] M. M. Moslehi und K. C. Saraswat " Studies of trapping and conduction in ultra-thin SiO_2 Gate Insulators" *in Technical Digest-IEEE 1984 International Electron Device Meeting*, vol. 30, pp. 157-160, 1984
- [120] E. Rosenbaum, Z. Liu, and C. Hu "Silicon dioxide breakdown lifetime enhancement under bipolar bias conditions" *IEEE Trans. Electron Devices*, vol. 40, no. 12, pp. 2287-2295, 1993
- [121] N. Nafria, J. Sune, D. Yelamos and X. Aymerich "High field dynamic stress of thin SiO_2 films" *Microelectron. Reliab.*, vol. 35, no. 3, pp. 539-553, 1995
- [122] N. Nafria, J. Sune, D. Yelamos and "Degradation and breakdown of thin silicon dioxide films under dynamic electrical stress" *IEEE Trans. Electron Dev.*, vol. 43, no. 12, pp. 2215-2226, 1996
- [123] R. Rodriques, N. Nafria, J. Sune and X. Aymerich " Analysis of the evolution of the trapped charge distribution in 10nm SiO_2 films during DC and bipolar dynamic stress “ *Microelectron. Reliab.*, vol. 37, no. 10/11, pp. 1517-1520, 1997
- [124] C. Buseret et al. "Quantitative study of charge trapping in SiO_2 during bipolar fowler-nordheim injection" *Journal of Non-Crystalline Solids*, vol. 322, pp. 191-198, 2003

- [125] Y. Liu, T. P. Chen, C.-H. Ang and S. Fung “ Power-law dependence of charge trapping on injected charge in very thin SiO_2 films“ *Jpn. Appl. Phys. Part 2*, vol. 41, no.4A, pp. L384-L386, 2002
- [126] A. El-Hdiy and Dj. Ziane “Electric bidirectional stress effects on metal-oxide-silicon capacitor“ *Appl. Phys.*, vol. 89, no. 2, 2001
- [127] Claudius Schimpf „Optimierung von Zuverlässigkeitsuntersuchungen, Prüfabläufen und Nacharbeitsprozessen in der Elektronikproduktion“ *Dissertation, Meisenbach Verlag Bamberg*, 2009
- [128] Automotive Electronic Council, „Failure Mechanism based stress test qualification for intergrated circuits“ *AEC-Q100-Rev-G*, 14 May 2007
- [129] W. Neher „Zuverlässigkeitsbetrachtungen an Aufbau- und Verbindungstechnologien für elektronische Steuergeräte im Kraftfahrzeug unter Hochtemperaturbeanspruchungen“ *Dissertation, Verlag Dr. A. Detert templin, Band 14*, 2006
- [130] D. Wilkie and M. Hensen “A study of EEPROM endurance correlation with wafer level reliability data“ *IEEE-IRW Final Report*, pp. 55-60, 1995
- [131] N. Baboux, C. Plossu, P. Boivin “A new technique to characterize endurance of EEPROM tunnel oxides “ *Journal of Non-Crystalline Solids*, vol. 351, pp. 1890-1896, 2005
- [132] S. -H. Soh, C. Messick, C. -C. Chen, C. -P. Liu, B. Baker, R. Peterson, L. Sadwick “Fast test at Wafer-Level for endurance of tunnel oxides“*IEEE-IRW Final Report*, 1995
- [133] L. Nebrich, “Entwicklung eines Markromodells für die Schaltungs- und Zuverlässigkeitssimulation von EEPROM-Zellen im erhöhten Temperaturbereich“ *Dissertation an der Christian-Albrechts-Universität zu Kiel*, 2002
- [134] C. Chen, W. L. Wilson and M. Smayling „Tunneling induced charge generation in SiO_2 thin films“ *J. Appl Phys.*, vol. 83, no. 7, pp. 3898-3905, 1998
- [135] R. Bouchakour, J. M. Mirabel, Ph. Boivin, O. Pizzuto „Modeling of a Floating-gate EEPROM Cell using a charge sheet Approach“ *IEEE Midwest Symp. on Circuits and Systems*, 2000